

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:31:31 ; Search time 28.7778 Seconds  
(without alignments)  
537.209 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_114\_150

Perfect score: 235

Sequence: 1 NCGSVPHDTWLPKCSLCKWHGQLRCFPOAFLPGCD 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgm2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgm2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 3: /cgm2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 4: /cgm2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 5: /cgm2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 6: /cgm2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	235	100.0	129	4	US-10-390-566-4
2	235	100.0	129	4	US-10-390-566-11
3	235	100.0	129	4	US-10-390-566-17
4	235	100.0	129	4	US-10-390-566-24
5	235	100.0	139	4	US-10-390-566-3
6	235	100.0	139	4	US-10-390-566-10
7	235	100.0	139	4	US-10-390-566-16
8	235	100.0	139	4	US-10-390-566-23
9	235	100.0	174	4	US-10-665-602-7
10	235	100.0	188	4	US-10-241-220-58
11	235	100.0	188	4	US-10-388-578-6
12	235	100.0	188	4	US-10-390-566-1
13	235	100.0	188	4	US-10-390-566-2
14	235	100.0	188	4	US-10-390-566-9
15	235	100.0	188	4	US-10-390-566-22
16	235	100.0	188	4	US-10-362-597A-3
17	235	100.0	188	4	US-10-362-597A-4
18	235	100.0	188	4	US-10-362-597A-6
19	235	100.0	188	4	US-10-362-597A-96
20	235	100.0	188	4	US-10-407-481-3
21	235	100.0	188	4	US-10-407-481-4
22	235	100.0	188	4	US-10-407-481-6
23	235	100.0	188	4	US-10-407-481-96
24	235	100.0	188	4	US-10-712-124-58
25	235	100.0	188	4	US-10-693-538-1
26	235	100.0	188	4	US-10-693-538-2
27	235	100.0	188	4	US-10-389-431-6

28	235	100.0	188	4	US-10-816-476-3
29	235	100.0	188	4	US-10-816-476-4
30	235	100.0	188	4	US-10-816-476-6
31	235	100.0	188	4	US-10-816-476-96
32	235	100.0	188	5	US-10-872-972-58
33	235	100.0	188	5	US-10-872-972-58
34	235	100.0	188	5	US-10-491-997-28
35	235	100.0	188	5	US-10-491-997-28
36	235	100.0	188	5	US-10-940-431-72
37	235	100.0	188	5	US-10-983-340-13
38	235	100.0	360	4	US-10-390-566-7
39	235	100.0	360	4	US-10-390-566-14
40	235	100.0	360	4	US-10-390-566-20
41	235	100.0	360	4	US-10-390-566-27
42	235	100.0	367	4	US-10-390-566-6
43	235	100.0	367	4	US-10-390-566-13
44	235	100.0	367	4	US-10-390-566-19
45	235	100.0	367	4	US-10-390-566-26

## ALIGNMENTS

RESULT 1  
US-10-390-566-4  
; Sequence 4, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 129  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-4  
Query Match 100.0%; Score 235; DB 4; Length 129;  
Best Local Similarity 100.0%; Pred. No. 2.8e-21;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 NCGSVPHDTWLPKCSLCKWHGQLRCFPOAFLPGCD 37  
|||||  
Db 74 NCGSVPHDTWLPKCSLCKWHGQLRCFPOAFLPGCD 110  
RESULT 2  
US-10-390-566-11  
; Sequence 11, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148

; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 129  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-11

Query Match 100.0%; Score 235; DB 4; Length 129;  
Best Local Similarity 100.0%; Pred. No. 2.8e-21;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37  
|||||  
DB 74 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 110

## RESULT 3

US-10-390-566-17  
; Sequence 17, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: Cripito Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 129  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-17

Query Match 100.0%; Score 235; DB 4; Length 129;  
Best Local Similarity 100.0%; Pred. No. 2.8e-21;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37  
|||||  
DB 74 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 110

## RESULT 4

US-10-390-566-24  
; Sequence 24, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: Cripito Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18

; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 129  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-24

Query Match 100.0%; Score 235; DB 4; Length 129;  
Best Local Similarity 100.0%; Pred. No. 2.8e-21;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37  
|||||  
DB 74 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 110

## RESULT 5

US-10-390-566-3  
; Sequence 3, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: Cripito Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-3

Query Match 100.0%; Score 235; DB 4; Length 139;  
Best Local Similarity 100.0%; Pred. No. 3e-21;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37  
|||||  
DB 84 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 120

## RESULT 6

US-10-390-566-10  
; Sequence 10, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: Cripito Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10

```
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-10

Query Match      100.0%; Score 235; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 3e-21;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKCKSLCKCKWHGQLRCFPPQAFPLPGCD 37
    |||||
Db 84 NCGSVPHDTWLPKCKSLCKCKWHGQLRCFPPQAFPLPGCD 120

RESULT 7
US-10-390-566-16
; Sequence 16, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Cripco Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR FILING DATE: 2003-03-17
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-16

Query Match      100.0%; Score 235; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 3e-21;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKCKSLCKCKWHGQLRCFPPQAFPLPGCD 37
    |||||
Db 84 NCGSVPHDTWLPKCKSLCKCKWHGQLRCFPPQAFPLPGCD 120

RESULT 8
US-10-390-566-23
; Sequence 23, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Cripco Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR FILING DATE: 2003-03-17
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo Sapien
```

```
US-10-390-566-23

Query Match      100.0%; Score 235; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 3e-21;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKCKSLCKCKWHGQLRCFPPQAFPLPGCD 37
    |||||
Db 84 NCGSVPHDTWLPKCKSLCKCKWHGQLRCFPPQAFPLPGCD 120

RESULT 9
US-10-665-602-7
; Sequence 7, Application US/10665602
; Publication No. US20040086967A1
; GENERAL INFORMATION:
; APPLICANT: Meisener, Paul S.
; APPLICANT: Coleman, Timothy A.
; TITLE OF INVENTION: Human Cripitin Growth Factor
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/665,602
; FILING DATE: 22-Sep-2003
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/393,023A
; FILING DATE: 09-SEP-1999
; APPLICATION NUMBER: US 08/471,371
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Marks, Michelle S.
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PP200D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-665-602-7

Query Match      100.0%; Score 235; DB 4; Length 174;
Best Local Similarity 100.0%; Pred. No. 3.7e-21;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKCKSLCKCKWHGQLRCFPPQAFPLPGCD 37
    |||||
Db 100 NCGSVPHDTWLPKCKSLCKCKWHGQLRCFPPQAFPLPGCD 136

RESULT 10
US-10-241-220-58
; Sequence 58, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
```

```
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Williams, W. Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 58
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-241-220-58

Query Match      100.0%; Score 235; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 4e-21;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NCGSVPHDTWLPKCKSLCKCKWHGQLRCPFPQAFPLPGCD 37
DB 114 NCGSVPHDTWLPKCKSLCKCKWHGQLRCPFPQAFPLPGCD 150

RESULT 11
US-10-388-578-6
; Sequence 6, Application US/10388578
; Publication No. US20030224411A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Stanton, Lawrence
; APPLICANT: Ralph, Brandenberger
; APPLICANT: Joseph, Gold D.
; APPLICANT: John, Irving
; APPLICANT: Mandalam, Ramkumar
; APPLICANT: Mok, Michael
; APPLICANT: Shelton, Dawne
; TITLE OF INVENTION: Genes that are Up- or Down-Regulated During Differentiation of Hu
; TITLE OF INVENTION: Embryonic Stem Cells
; FILE REFERENCE: 135/001
; CURRENT APPLICATION NUMBER: US/10/388,578
; CURRENT FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: Custom
; SEQ ID NO 6
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-388-578-6

Query Match      100.0%; Score 235; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 4e-21;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NCGSVPHDTWLPKCKSLCKCKWHGQLRCPFPQAFPLPGCD 37
DB 114 NCGSVPHDTWLPKCKSLCKCKWHGQLRCPFPQAFPLPGCD 150

RESULT 12
US-10-390-566-1
; Sequence 1, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
```

```
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-1

Query Match      100.0%; Score 235; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 4e-21;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NCGSVPHDTWLPKCKSLCKCKWHGQLRCPFPQAFPLPGCD 37
DB 114 NCGSVPHDTWLPKCKSLCKCKWHGQLRCPFPQAFPLPGCD 150

RESULT 13
US-10-390-566-2
; Sequence 2, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Foley, Susan
; APPLICANT: Williams, Kevin P.
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-2

Query Match      100.0%; Score 235; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 4e-21;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NCGSVPHDTWLPKCKSLCKCKWHGQLRCPFPQAFPLPGCD 37
DB 114 NCGSVPHDTWLPKCKSLCKCKWHGQLRCPFPQAFPLPGCD 150

RESULT 14
US-10-390-566-9
; Sequence 9, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
```



; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-9

Query Match 100.0%; Score 235; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 4e-21;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37  
Db 114 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 150

## RESULT 15

US-10-390-566-22  
; Sequence 22, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Santicola-Nadel, Michele  
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; PRIOR FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-22

Query Match 100.0%; Score 235; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 4e-21;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37  
Db 114 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 150

Search completed: February 3, 2006, 18:39:09  
Job time : 29.7778 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:33:31 ; Search time 2.13778 Seconds  
(without alignments)  
202.813 Million cell updates/sec

Title: us-10-693-538-1\_COPY\_114\_150

Perfect score: 235

Sequence: 1 NCGSVPHDTWLPKCSLCKWHGQLRCFPOAFLPGCD 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	235	100.0	188	6	US-10-945-853-1
2	235	100.0	188	6	US-10-945-853-2
3	201	85.5	32	6	US-10-945-853-9
4	97	41.3	15	6	US-10-945-853-6
5	87	37.0	28	6	US-10-945-853-7
6	86	36.6	1477	7	US-11-149-003-8
7	86	36.6	1512	7	US-11-149-003-10
8	81	34.5	13	6	US-10-945-853-5
9	78	33.2	1535	7	US-11-149-003-14
10	78	33.2	1570	7	US-11-149-003-12
11	78	33.2	1593	7	US-11-149-003-4
12	78	33.2	1628	7	US-11-149-003-2
13	67	28.5	1057	7	US-11-149-003-6
14	67	28.5	1192	7	US-11-149-003-18
15	67	28.5	1207	7	US-11-149-003-20
16	67	28.5	1251	7	US-11-149-003-16
17	67	28.5	1342	7	US-11-149-003-24
18	64.5	27.4	308	7	US-11-043-788-390
19	64.5	27.4	344	7	US-11-043-788-397
20	64.5	27.4	356	7	US-11-043-788-472
21	64.5	27.4	366	7	US-11-043-788-396
22	64.5	27.4	378	7	US-11-043-788-471
23	64.5	27.4	386	7	US-11-043-788-398
24	64.5	27.4	388	7	US-11-043-788-465
25	64.5	27.4	392	7	US-11-043-788-469

26 64.5 27.4 398 7 US-11-043-788-463  
27 64.5 27.4 429 7 US-11-043-788-389  
28 64.5 27.4 431 7 US-11-043-788-395  
29 64.5 27.4 439 7 US-11-043-788-466  
30 64.5 27.4 443 7 US-11-043-788-470  
31 64.5 27.4 451 6 US-10-083-703-82  
32 64.5 27.4 451 7 US-11-102-240-82  
33 64.5 27.4 451 7 US-11-043-788-394  
34 64.5 27.4 539 7 US-11-043-788-464  
35 64.5 27.4 549 7 US-11-043-788-467  
36 64.5 27.4 549 7 US-11-043-788-468  
37 62.5 26.6 816 7 US-11-050-439-48  
38 60 25.5 1464 7 US-11-000-463-243  
39 60 25.5 1464 7 US-11-186-284-28  
40 60 25.5 1464 7 US-11-021-603-2  
41 60 25.5 1467 6 US-10-821-234-1096  
42 58 24.7 1166 6 US-10-821-234-964  
43 58 24.7 1466 7 US-11-186-284-33  
44 57 24.3 1496 7 US-11-186-284-35  
45 56.5 24.0 1018 7 US-11-067-121-17

## ALIGNMENTS

RESULT 1  
US-10-945-853-1  
; Sequence 1, Application US/10945853  
; Publication No. US20050255117A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc.  
; APPLICANT: Sanicola-Nadel, Michele  
; APPLICANT: Adkins, Heather  
; APPLICANT: Miklaez, Steven Donald  
; APPLICANT: Rayhorn, Paul  
; APPLICANT: Schiffer, Susan Gail  
; APPLICANT: Williams, Kevin  
; TITLE OF INVENTION: Crypto-Specific Antibodies  
; FILE REFERENCE: BGN117GPPCCN  
; CURRENT APPLICATION NUMBER: US/10/945,853  
; CURRENT FILING DATE: 2004-09-20  
; PRIOR APPLICATION NUMBER: PCT/US02/31462  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: PCT/US02/11950  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/286,782  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: 60/293,020  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/301,091  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/367,002  
; PRIOR FILING DATE: 2002-03-22  
; NUMBER OF SEQ ID NOS: 9  
; SEQ ID NO 1  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-945-853-1

Query Match 100.0%; Score 235; DB 6; Length 188;  
Best Local Similarity 100.0%; Pred. No. 5.5e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKCSLCKWHGQLRCFPOAFLPGCD 37  
Db 114 NCGSVPHDTWLPKCSLCKWHGQLRCFPOAFLPGCD 150

RESULT 2  
US-10-945-853-2  
; Sequence 2, Application US/10945853  
; Publication No. US20050255117A1

```

; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Sanicola-Nadel, Michele
; APPLICANT: Adkins, Heather
; APPLICANT: Miklasz, Steven Donald
; APPLICANT: Rayhorn, Paul
; APPLICANT: Schiffer, Susan Gail
; APPLICANT: Williams, Kevin
; TITLE OF INVENTION: CripTo-Specific Antibodies
; FILE REFERENCE: BGN117CPPCCN
; CURRENT APPLICATION NUMBER: US/10/945,853
; CURRENT FILING DATE: 2004-09-20
; PRIOR APPLICATION NUMBER: PCT/US02/31462
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: PCT/US02/11950
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/286,782
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/293,020
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/301,091
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/367,002
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 2
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-945-853-2

Query Match 100.0%; Score 235; DB 6; Length 188;
Best Local Similarity 100.0%; Pred. No. 5.5e-23;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37
Db 114 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 150

RESULT 3
; Sequence 9, Application US/10945853
; Publication No. US20050255117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Sanicola-Nadel, Michele
; APPLICANT: Adkins, Heather
; APPLICANT: Miklasz, Steven Donald
; APPLICANT: Rayhorn, Paul
; APPLICANT: Schiffer, Susan Gail
; APPLICANT: Williams, Kevin
; TITLE OF INVENTION: CripTo-Specific Antibodies
; FILE REFERENCE: BGN117CPPCCN
; CURRENT APPLICATION NUMBER: US/10/945,853
; CURRENT FILING DATE: 2004-09-20
; PRIOR APPLICATION NUMBER: PCT/US02/31462
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: PCT/US02/11950
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/286,782
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/293,020
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/301,091
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/367,002
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 9
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; NAME/KEY: Mod_res
; LOCATION: 1
; OTHER INFORMATION: N-terminal acetylation
US-10-945-853-9

Query Match 85.5%; Score 201; DB 6; Length 32;
Best Local Similarity 96.9%; Pred. No. 2.5e-19;
Matches 31; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 PHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37
Db 1 PHNTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 32

RESULT 4
US-10-945-853-6
; Sequence 6, Application US/10945853
; Publication No. US20050255117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Sanicola-Nadel, Michele
; APPLICANT: Adkins, Heather
; APPLICANT: Miklasz, Steven Donald
; APPLICANT: Rayhorn, Paul
; APPLICANT: Schiffer, Susan Gail
; APPLICANT: Williams, Kevin
; TITLE OF INVENTION: CripTo-Specific Antibodies
; FILE REFERENCE: BGN117CPPCCN
; CURRENT APPLICATION NUMBER: US/10/945,853
; CURRENT FILING DATE: 2004-09-20
; PRIOR APPLICATION NUMBER: PCT/US02/31462
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: PCT/US02/11950
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/286,782
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/293,020
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/301,091
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/367,002
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 6
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-945-853-6

Query Match 41.3%; Score 97; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 SLCKWHGQLRCFPQ 30
Db 1 SLCKWHGQLRCFPQ 15

RESULT 5
US-10-945-853-7
; Sequence 7, Application US/10945853
; Publication No. US20050255117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Sanicola-Nadel, Michele
; APPLICANT: Adkins, Heather
; APPLICANT: Miklasz, Steven Donald
; APPLICANT: Rayhorn, Paul
; APPLICANT: Schiffer, Susan Gail
; APPLICANT: Williams, Kevin
; TITLE OF INVENTION: CripTo-Specific Antibodies
; FILE REFERENCE: BGN117CPPCCN

```

; CURRENT APPLICATION NUMBER: US/10/945,853  
; CURRENT FILING DATE: 2004-09-20  
; PRIOR APPLICATION NUMBER: PCT/US02/31462  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: PCT/US02/11950  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/286,782  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: 60/293,020  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/301,091  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/367,002  
; PRIOR FILING DATE: 2002-03-22  
; NUMBER OF SEQ ID NOS: 9  
; SEQ ID NO 7  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: Mod\_res  
; LOCATION: 1  
; OTHER INFORMATION: N-terminal acetylation  
US-10-945-853-7

Query Match 37.0%; Score 87; DB 6; Length 28;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKC 14  
| | | | | | | | | | | | | | | | | |  
Db 15 NCGSVPHDTWLPKC 28

RESULT 6  
US-11-149-003-8  
; Sequence 8, Application US/11149003  
; Publication No. US20060014277A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; APPLICANT: Turner, C. Alexander Jr.  
; TITLE OF INVENTION: Novel Human Kiellin-like Proteins and Polynucleotides Encoding the  
; FILE REFERENCE: LEX-0360-USA  
; CURRENT APPLICATION NUMBER: US/11/149,003  
; CURRENT FILING DATE: 2005-06-09  
; PRIOR APPLICATION NUMBER: US/10/189,971  
; PRIOR FILING DATE: 2002-07-03  
; PRIOR APPLICATION NUMBER: US 60/302,949  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: US 60/315,634  
; PRIOR FILING DATE: 2001-08-29  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 1477  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-149-003-8

Query Match 36.6%; Score 86; DB 7; Length 1477;  
Best Local Similarity 51.9%; Pred. No. 0.0014;  
Matches 14; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy 10 WLPKCSLCKWQHLCRCFQAFLPGC 36  
| | | | | | | | | | | | | | | | | |  
Db 109 WEPDCTACVCQDGAHCGFQAHLPGC 135

RESULT 7  
US-11-149-003-10  
; Sequence 10, Application US/11149003

; Publication No. US20060014277A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; APPLICANT: Turner, C. Alexander Jr.  
; TITLE OF INVENTION: Novel Human Kiellin-like Proteins and Polynucleotides Encoding the  
; FILE REFERENCE: LEX-0360-USA  
; CURRENT APPLICATION NUMBER: US/11/149,003  
; CURRENT FILING DATE: 2005-06-09  
; PRIOR APPLICATION NUMBER: US/10/189,971  
; PRIOR FILING DATE: 2002-07-03  
; PRIOR APPLICATION NUMBER: US 60/302,949  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: US 60/315,634  
; PRIOR FILING DATE: 2001-08-29  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 1512  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-149-003-10

Query Match 36.6%; Score 86; DB 7; Length 1512;  
Best Local Similarity 51.9%; Pred. No. 0.0015;  
Matches 14; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy 10 WLPKCSLCKWQHLCRCFQAFLPGC 36  
| | | | | | | | | | | | | | | | | |  
Db 109 WEPDCTACVCQDGAHCGFQAHLPGC 135

RESULT 8  
US-10-945-853-5  
; Sequence 5, Application US/10945853  
; Publication No. US20050255117A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc.  
; APPLICANT: Sanicola-Nadel, Michele  
; APPLICANT: Atkins, Heather  
; APPLICANT: Miklasz, Steven Donald  
; APPLICANT: Rayhorn, Paul  
; APPLICANT: Schiffer, Susan Gail  
; APPLICANT: Williams, Kevin  
; TITLE OF INVENTION: Crypto-Specific Antibodies  
; FILE REFERENCE: BGN117CPCCN  
; CURRENT APPLICATION NUMBER: US/10/945,853  
; CURRENT FILING DATE: 2004-09-20  
; PRIOR APPLICATION NUMBER: PCT/US02/31462  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: PCT/US02/11950  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/286,782  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: 60/293,020  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/301,091  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/367,002  
; PRIOR FILING DATE: 2002-03-22  
; NUMBER OF SEQ ID NOS: 9  
; SEQ ID NO 5  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-945-853-5

Query Match 34.5%; Score 81; DB 6; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GSVPHDTWLPKC 15

db 1 GSVPHDTWLPKKC 13

```

RESULT 9
US-11-149-003-14
; Sequence 14, Application US/11149003
; Publication No. US20060014277A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: Novel Human Kiein-li
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0360-USA
; CURRENT APPLICATION NUMBER: US/11/149,003
; CURRENT FILING DATE: 2005-06-09
; PRIOR APPLICATION NUMBER: US/10/189,971
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/302,949
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/315,634
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1535
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-149-003-14

```

```

Query Match      33.2%; Score 78; DB 7; Length 1535;
Best Local Similarity 48.1%; Pred.No. 0.015;
Matches 13; Conservative 1; Mismatches 13; Indels

Qy      10 WLPKKCSLCKWHGQLRCPQAFPGC 36
      | : | | | | | | | | | |
Db      42 WEPDACTACVCODGAAGCGPOAHLPHC 68

```

```

RESULT 10
US-11-149-003-12
, Sequence 12, Application US/11149003
, Publication No. US20060014277A1
, GENERAL INFORMATION:
, APPLICANT: Walke, D. Wade
, APPLICANT: Scoville, John
, APPLICANT: Turner, C. Alexander Jr.
, TITLE OF INVENTION: Novel Human Kielin-li-
, TITLE OF INVENTION: Same
, FILE REFERENCE: LEX-0360-USA
, CURRENT APPLICATION NUMBER: US/11/149,003
, CURRENT FILING DATE: 2005-06-09
, PRIOR APPLICATION NUMBER: US/10/189,971
, PRIOR FILING DATE: 2002-07-03
, PRIOR APPLICATION NUMBER: US 60/302,949
, PRIOR FILING DATE: 2001-07-03
, PRIOR APPLICATION NUMBER: US 60/315,634
, PRIOR FILING DATE: 2001-08-29
, NUMBER OF SEQ ID NOS: 25
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 12
, LENGTH: 1570
, TYPE: PRT
, ORGANISM: homo sapiens
US-11-149-003-12

```

Query Match 33.2%; Score 78;  
Best Local Similarity 48.1%; Pred. NO.  
Matches 13; Conservative 1; Mismatch

QY 10 WLPKKCSLCKWHGQLRCPQAFLPQC 36

Db 42 WEPDACTACVCDGAHCGPQAHLPHC 68

```

RESULT 11
US-11-149-003-4
; Sequence 4, Application US/11149003
; Publication No. US20060014277A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: Novel Human Kiellin-li
; FILE REFERENCE: Same
; CURRENT APPLICATION NUMBER: US/11/149,003
; CURRENT FILING DATE: 2005-06-09
; PRIOR APPLICATION NUMBER: US/10/189,971
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/302,949
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/315,634
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1593
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-149-003-4

```

Query Match	33.2%	Score 78;	DB 7;	Length 1593;
Best Local Similarity	48.1%;	Pred. No. 0.015;		
Matches 13;	Conservative 1;	Mismatches 13;	Indels 0;	Gaps 0;
QY	10	WLPKKSLCKWHGQLRCFFQAFLPGC	36	
DP	109	WEPDACTACVCGDGAAHCGFOAHLPHC	135	

```

RESULT 12
US-11-149-003-2
; Sequence 2, Application US/11149003
; Publication No. US20060014277A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: Novel Human Kiefin-li
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0360-USA
; CURRENT APPLICATION NUMBER: US/11/149,003
; CURRENT FILING DATE: 2005-06-09
; PRIOR APPLICATION NUMBER: US/10/189,971
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/302,949
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/315,634
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0.
; SEQ ID NO 2
; LENGTH: 1628
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-149-003-2

```

	Query Match	33.2%	Score 78	DB 7	Length 1628
	Best Local Similarity	48.1%	Pred. No.	0.016	
	Matches 13	Conservative	1	Mismatches 13	Indels 0
Qy	10	WLPKKCSLCKWHGQLRCFPOAFLPGC	36		
pb	109	WEEDACTACVCGDGAHHCPOAHLPHC	135		

RESULT 13  
US-11-149-003-6  
; Sequence 6, Application US/11149003  
; Publication No. US20060014277A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; APPLICANT: Turner, C. Alexander Jr.  
; TITLE OF INVENTION: Novel Human Kiellin-like Proteins and Polynucleotides Encoding the  
; FILE REFERENCE: LEX-0360-USA  
; CURRENT APPLICATION NUMBER: US/11/149,003  
; CURRENT FILING DATE: 2005-06-09  
; PRIOR APPLICATION NUMBER: US/10/189,971  
; PRIOR FILING DATE: 2002-07-03  
; PRIOR APPLICATION NUMBER: US 60/302,949  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: US 60/315,634  
; PRIOR FILING DATE: 2001-08-29  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1057  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-11-149-003-6

Query Match 28.5%; Score 67; DB 7; Length 1057;  
Best Local Similarity 50.0%; Pred. No. 0.27;  
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 15 CSLCKWHGOLRCFPQAFPLP 34  
Db 172 CHLCLWEGSVSCEPKACAP 191

RESULT 14  
US-11-149-003-18  
; Sequence 18, Application US/11149003  
; Publication No. US20060014277A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; APPLICANT: Turner, C. Alexander Jr.  
; TITLE OF INVENTION: Novel Human Kiellin-like Proteins and Polynucleotides Encoding the  
; FILE REFERENCE: LEX-0360-USA  
; CURRENT APPLICATION NUMBER: US/11/149,003  
; CURRENT FILING DATE: 2005-06-09  
; PRIOR APPLICATION NUMBER: US/10/189,971  
; PRIOR FILING DATE: 2002-07-03  
; PRIOR APPLICATION NUMBER: US 60/302,949  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: US 60/315,634  
; PRIOR FILING DATE: 2001-08-29  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 1192  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-11-149-003-18

Query Match 28.5%; Score 67; DB 7; Length 1192;  
Best Local Similarity 50.0%; Pred. No. 0.3;  
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 15 CSLCKWHGOLRCFPQAFPLP 34  
Db 342 CHLCLWEGSVSCEPKACAP 361

RESULT 15  
US-11-149-003-20  
; Sequence 20, Application US/11149003  
; Publication No. US20060014277A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; APPLICANT: Turner, C. Alexander Jr.  
; TITLE OF INVENTION: Novel Human Kiellin-like Proteins and Polynucleotides Encoding the  
; FILE REFERENCE: LEX-0360-USA  
; CURRENT APPLICATION NUMBER: US/11/149,003  
; CURRENT FILING DATE: 2005-06-09  
; PRIOR APPLICATION NUMBER: US/10/189,971  
; PRIOR FILING DATE: 2002-07-03  
; PRIOR APPLICATION NUMBER: US 60/302,949  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: US 60/315,634  
; PRIOR FILING DATE: 2001-08-29  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 1207  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-11-149-003-20

Query Match 28.5%; Score 67; DB 7; Length 1207;  
Best Local Similarity 50.0%; Pred. No. 0.3;  
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 15 CSLCKWHGOLRCFPQAFPLP 34  
Db 357 CHLCLWEGSVSCEPKACAP 376

Search completed: February 3, 2006, 18:39:40  
Job time : 2.13778 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 14:27:36 ; Search time 139 Seconds  
(without alignments)  
110.635 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_77\_111  
Perfect score: 211  
Sequence: 1 ELNRTCLNGTCLMLSGFCACPPSPYGRNCEHVR 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211	100.0	38	5	Abb77105 Human Cri
2	211	100.0	129	5	Abb77104 Human Cri
3	211	100.0	139	5	Abb77103 Human Cri
4	211	100.0	173	8	Ad88698 Amino aci
5	211	100.0	174	2	AAR13326
6	211	100.0	174	2	AAW32107
7	211	100.0	174	8	Ad005065 Human cri
8	211	100.0	183	8	Ad88699 Amino aci
9	211	100.0	188	2	AAR22548
10	211	100.0	188	2	AAW29735
11	211	100.0	188	2	AAW87630
12	211	100.0	188	5	Abb77101 Human Cri
13	211	100.0	188	5	AAO14638
14	211	100.0	188	5	AAO14636
15	211	100.0	188	5	AAO14727
16	211	100.0	188	6	ABP97176
17	211	100.0	188	6	ABP58131
18	211	100.0	188	7	ADC78799
19	211	100.0	188	7	ADD26058 CR-1 cri
20	211	100.0	188	8	Ad182176 Human ter
21	211	100.0	188	8	Ad70486 Human ter
22	211	100.0	188	8	Ad88697 Amino aci
23	211	100.0	188	8	Ad79200 Human Cri
24	211	100.0	188	8	Ad79197 Human Cri

25	211	100.0	188	8	ADT79290
26	211	100.0	188	9	ADY85961 Human Cri
27	211	100.0	188	9	ADZ42238 Human Cri
28	211	100.0	266	5	Abb77108 Human Cri
29	211	100.0	360	5	Abb77107 Human Cri
30	211	100.0	367	5	Abb77106 Human Cri
31	206	97.6	38	5	Abb77112 Human Cri
32	206	97.6	129	5	Abb77111 Human Cri
33	206	97.6	139	5	Abb77110 Human Cri
34	206	97.6	188	5	Abb77109 Human Cri
35	206	97.6	266	5	Abb77115 Human Cri
36	206	97.6	360	5	Abb77114 Human Cri
37	206	97.6	367	5	Abb77113 Human Cri
38	204	96.7	37	6	ABP54994
39	203	96.2	38	5	Abb77118 Human Cri
40	203	96.2	129	5	Abb77117 Human Cri
41	203	96.2	139	5	Abb77123 Human Cri
42	203	96.2	139	5	Abb77116 Human Cri
43	203	96.2	188	2	AAW22547
44	203	96.2	188	2	AAW25667
45	203	96.2	188	2	AAW19980

ALIGNMENTS

RESULT 1  
ID ABB77105 standard; protein; 38 AA.  
XX ABB77105;  
AC ABB77105;  
XX  
DT 08-OCT-2002 (first entry)  
XX  
DE Human Cripto-1 fragment (CR(75-112)).  
XX  
KW Human; Cripto-1; CR-1; mutant; tumour; cell proliferation.  
XX  
OS Homo sapiens.  
XX  
PN WO200222808-A2.  
PD 21-MAR-2002.  
XX  
PF 18-SEP-2001; 2001WO-US029066.  
XX  
PR 18-SEP-2000; 2000US-0233148P.  
XX  
PA (BIOJ ) BIOGEN INC.  
XX  
PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
XX  
DR WPI; 2002-339868/37.  
XX  
PT New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful  
for treating cell proliferation, especially cancer, comprises amino acid  
change that prevents fucosylation at Thr88.  
XX  
PS Claim 2; Page 35; 41pp; English.  
XX  
CC The sequence represents a fragment of the human Cripto-1 polypeptide. The  
invention relates to a novel mutant CRIPTO polypeptide, or its functional  
fragment, having at least one amino acid alteration at positions 86, 87  
or 88. The mutant polypeptide, or its chimera, is used to inhibit growth  
of tumour cells, in vivo or in vitro, particularly for treating breast,  
ovarian, renal, colorectal, uterine, prostatic, lung, bladder or central  
nervous system cancers, melanoma and leukaemia, also generally for  
treating undesired cell proliferation  
XX  
SQ Sequence 38 AA;

Query Match 100.0%; Score 211; DB 5; Length 38;  
Best Local Similarity 100.0%; Pred. No. 7.7e-14;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCCLLGGTCMLGSCFACPPSPFYGRNCEHDVR 35  
 DB 3 ELNRTCCLLGGTCMLGSCFACPPSPFYGRNCEHDVR 37

RESULT 2  
 ABB77104  
 ID ABB77104 standard; protein; 129 AA.  
 AC ABB77104;  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Human Cripto-1 C-terminal truncated form (CR38(cc)).  
 KW Human; Cripto-1; CR-1; mutant; tumour; cell proliferation.  
 XX Homo sapiens.  
 OS  
 PN WO200222808-A2.  
 XX  
 PD 21-MAR-2002.  
 XX  
 PF 18-SEP-2001; 2001WO-US029066.  
 XX  
 PR 18-SEP-2000; 2000US-0233148P.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
 XX WPI; 2002-339868/37.  
 DR  
 XX  
 PT New mutant form of CRYPTO (teratocarcinoma-derived growth factor), useful  
 PT for treating cell proliferation, especially cancer, comprises amino acid  
 PT change that prevents fucosylation at Thr88.  
 XX  
 PS Claim 2; Page 35; 41pp; English.  
 CC  
 CC The sequence represents a C-terminal truncated form of human Cripto-1.  
 CC The invention relates to a novel mutant CRYPTO polypeptide, or its  
 CC functional fragment, having at least one amino acid alteration at  
 CC positions 86, 87 or 88. The mutant polypeptide, or its chimera, is used  
 CC to inhibit growth of tumour cells, in vivo or in vitro, particularly for  
 CC treating breast, ovarian, renal, colorectal, uterine, prostatic, lung,  
 CC bladder or central nervous system cancers, melanoma and leukaemia, also  
 CC generally for treating undesired cell proliferation  
 XX  
 SQ Sequence 129 AA;

Query Match 100.0%; Score 211; DB 5; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 2e-13;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCCLLGGTCMLGSCFACPPSPFYGRNCEHDVR 35  
 DB 37 ELNRTCCLLGGTCMLGSCFACPPSPFYGRNCEHDVR 71

RESULT 3  
 ABB77103  
 ID ABB77103 standard; protein; 139 AA.  
 AC ABB77103;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Human Cripto-1 C-terminal truncated form (CR(cc)).  
 KW Human; Cripto-1; CR-1; mutant; tumour; cell proliferation.  
 XX

OS Homo sapiens.  
 XX  
 PN WO200222808-A2.  
 XX  
 PD 21-MAR-2002.  
 XX  
 PF 18-SEP-2001; 2001WO-US029066.  
 XX  
 PR 18-SEP-2000; 2000US-0233148P.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
 XX WPI; 2002-339868/37.  
 DR  
 XX  
 PT New mutant form of CRYPTO (teratocarcinoma-derived growth factor), useful  
 PT for treating cell proliferation, especially cancer, comprises amino acid  
 PT change that prevents fucosylation at Thr88.  
 XX  
 PS Claim 2; Page 35; 41pp; English.  
 CC  
 CC The sequence represents a C-terminal truncated form of human Cripto-1.  
 CC The invention relates to a novel mutant CRYPTO polypeptide, or its  
 CC functional fragment, having at least one amino acid alteration at  
 CC positions 86, 87 or 88. The mutant polypeptide, or its chimera, is used  
 CC to inhibit growth of tumour cells, in vivo or in vitro, particularly for  
 CC treating breast, ovarian, renal, colorectal, uterine, prostatic, lung,  
 CC bladder or central nervous system cancers, melanoma and leukaemia, also  
 CC generally for treating undesired cell proliferation  
 XX  
 SQ Sequence 139 AA;

Query Match 100.0%; Score 211; DB 5; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-13;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCCLLGGTCMLGSCFACPPSPFYGRNCEHDVR 35  
 DB 47 ELNRTCCLLGGTCMLGSCFACPPSPFYGRNCEHDVR 81

RESULT 4  
 ADS88698  
 ID ADS88698 standard; protein; 173 AA.  
 XX  
 AC ADS88698;  
 XX  
 DT 16-DEC-2004 (first entry)  
 XX  
 DE Amino acid sequence of a human secreted Cripto protein.  
 XX  
 KW stem cell differentiation; cardiomyocyte; stem cell;  
 KW epidermal growth factor; EGF; EGF-CFC family; Cripto; neuronal cell;  
 KW Alq4; Nodal; heart disease; neuropathology; myocardial infarction;  
 KW Parkinson's disease; Alzheimer's disease; retinal degeneration;  
 KW Cell therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004083375-A2.  
 XX  
 PD 30-SEP-2004.  
 XX  
 PF 19-MAR-2004; 2004WO-IT000133.  
 XX  
 PR 21-MAR-2003; 2003IT-RM000125.  
 XX  
 PR 29-JUL-2003; 2003IT-RM000370.  
 XX  
 PA (MINC/) MINCHIOTTI G.  
 PA (PERS/) PERSICO M.  
 PA (PARI/) PARISI S.  
 XX

PI Minchiotti G, Persico M, Parisi S;  
 DR WPI; 2004-691035/67.  
 XX  
 XX Inducing stem cell differentiation in cardiomyocytes or neuronal cells by  
 PT exposure to a protein of EGF-CFC or Cripto protein inhibitors,  
 PT respectively, useful in treating heart diseases and neuropathologies.  
 XX  
 XX Disclosure; Page 13; 35pp; English.  
 PS  
 CC The specification describes a method for inducing stem cell  
 CC differentiation into cardiomyocytes. The method comprises exposing the  
 CC cells to a protein of the epidermal growth factor (EGF)-CFC family or its  
 CC derivatives and having at least the EGF and CFC domains (e.g. Cripto  
 CC protein). A second method is described for stem cell differentiation into  
 CC neuronal cells, where the cells are exposed to an inhibitor of the Cripto  
 CC protein. The Cripto protein inhibitor is an anti-Cripto antibody or its  
 CC functional fragments, or is a peptide specifically selected from a random  
 CC combinatorial peptide library, or is an antagonist of the Alq4(receptor)-  
 CC Cripto(co-receptor)-Nodal(ligand) pathway. The differentiated stem cells  
 CC of the invention are useful in the treatment of heart diseases and  
 CC neuropathologies, e.g. myocardial infarction, Parkinson's disease,  
 CC Alzheimer's disease and retinal degeneration. They are especially useful  
 CC for cell therapy. The Cripto protein or its inhibitors is useful in the  
 CC preparation of a composition able to direct stem cell differentiation  
 CC toward the neuronal lineage. The present sequence represents a Cripto  
 CC protein, which maybe used in the method of the invention.  
 XX  
 XX Sequence 173 AA;  
 SQ

Query Match 100.0%; Score 211; DB 8; Length 173;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-13;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCCLLGGTCTMLGSCFCACPPSPFYGRNCEHDVR 35  
 |||||  
 DB 77 ELNRTCCLLGGTCTMLGSCFCACPPSPFYGRNCEHDVR 111  
 |||||

RESULT 5  
 AAR13326  
 ID AAR13326 standard; protein; 174 AA.  
 AC AAR13326;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 17-DEC-2001 (revised)  
 DT 22-OCT-1991 (first entry)  
 XX  
 XX Recombinant CRIPTO protein.  
 DE Cancer.  
 KW  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX USN7530165-N.  
 PN  
 XX  
 XX 09-JUL-1991.  
 PD  
 XX  
 XX 29-MAY-1990; 90US-00143529.  
 PF  
 XX  
 XX 29-MAY-1990; 90US-00530165.  
 PR  
 XX  
 XX (USSH ) NAT INST OF HEALTH.  
 PA  
 XX  
 XX WPI; 1991-245692/33.  
 DR  
 XX  
 XX New human CRIPTO gene - used for developing prods. for the study,  
 PT diagnosis prognosis and treatment of human cancers.  
 PT  
 XX  
 XX Disclosure; Fig 2; 24pp; English.  
 PS  
 XX  
 XX The sequence given is that of the recombinant E.coli derived CRIPTO

CC protein, produced from CRIPTO cDNA. The availability of this protein make  
 CC it possible to detect cells expressing the CRIPTO gene. It can be used to  
 CC develop oligonucleotide probes, antisense oligonucleotides and antibodies  
 CC for the study, diagnosis, prognosis and treatment of human cancer. The  
 CC cDNA sequence has been deposited in the EMBL Gene Data Bank. This  
 CC sequence and the amino acid sequence of the corresponding natural CRIPTO  
 CC protein (different to the recombinant sequence) are given in the  
 CC specification but poor print quality prevents their inclusion in this  
 CC database. (Note: Revised entry submitted to correct the patent number  
 CC format of US Government-owned NTIS applications to prevent clashes with  
 CC ongoing US granted patent numbers. For further information please visit  
 CC the Derwent web site at [www.derwent.com/dwpi/updates/ntis\\_us.html](http://www.derwent.com/dwpi/updates/ntis_us.html).)  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 XX  
 XX Sequence 174 AA;  
 SQ

Query Match 100.0%; Score 211; DB 2; Length 174;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-13;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCCLLGGTCTMLGSCFCACPPSPFYGRNCEHDVR 35  
 |||||  
 DB 63 ELNRTCCLLGGTCTMLGSCFCACPPSPFYGRNCEHDVR 97  
 |||||

RESULT 6  
 AAW32107  
 ID AAW32107 standard; protein; 174 AA.  
 AC AAW32107;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 17-FEB-1998 (first entry)  
 XX  
 XX Recombinant human CRIPTO protein expressed in E. coli.  
 DE  
 XX  
 XX CRIPTO gene; human; colorectal carcinoma; breast carcinoma; cancer;  
 KW tumour; diagnosis; prognosis; therapy; antibody; Escherichia coli.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX Synthetic.  
 OS  
 XX  
 XX US5654140-A.  
 PN  
 XX  
 XX 05-AUG-1997.  
 PD  
 XX  
 XX 10-NOV-1994; 94US-00337911.  
 PF  
 XX  
 XX 29-MAY-1990; 90US-00530165.  
 PR  
 XX  
 XX 18-SEP-1992; 92US-00947315.  
 PR  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX  
 XX Salomon DS, Persico MG;  
 PI  
 XX  
 XX WPI; 1997-401840/37.  
 DR  
 XX  
 XX New cloned CRIPTO gene - used for diagnosis and prognosis of carcinoma,  
 PT also anti-sense sequences for tumour treatment.  
 PT  
 XX  
 XX Disclosure; Fig 2; 13pp; English.  
 PS  
 XX  
 XX This protein comprises recombinant human CRIPTO produced in Escherichia  
 CC coli. Human CRIPTO cDNA (see AAT8973) was expressed in E. coli inclusion  
 CC bodies as a 20-22 kDa protein. CRIPTO is a member of the epidermal growth  
 CC factor family. Elevated levels of CRIPTO mRNA, or of protein translated  
 CC from it, are diagnostic and prognostic of carcinoma, especially  
 CC colorectal and breast carcinoma. Recombinant CRIPTO can be expressed in  
 CC prokaryotic or eukaryotic host cells, and used for the identification and  
 CC characterisation of specific cell surface receptors, or to raise specific  
 CC antibodies to assay CRIPTO protein expression in immunoassays. (Updated  
 CC on 25-MAR-2003 to correct PF field.)  
 XX

```
SQ Sequence 174 AA;
  Query Match      100.0%; Score 211; DB 2; Length 174;
  Best Local Similarity 100.0%; Pred. No. 2.5e-13;
  Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ELNRTCCCLNGGTCMLGSCFACACPPSPFYGRNCEHDVR 35
  |||||
DB 63 ELNRTCCCLNGGTCMLGSCFACACPPSPFYGRNCEHDVR 97

RESULT 7
ADO05065
ID ADO05065 standard; protein; 174 AA.
XX
AC ADO05065;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human cripto growth factor protein.
XX
KW Criptin growth factor; CGF; wound healing; tissue regeneration;
KW implant fixation; angiogenesis; neoplasia; tumour; gene therapy; human;
KW cripto growth factor.
XX
OS Homo sapiens.
XX
PN US2004086967-A1.
XX
PD 06-MAY-2004.
XX
PF 22-SEP-2003; 2003US-00665602.
XX
PR 06-JUN-1995; 95US-00471371.
PR 09-SEP-1999; 99US-00393023.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Meissner PS, Coleman TA;
XX
DR WPI; 2004-356201/33.
XX
PT New human polynucleotides encoding human criptin growth factor
PT polypeptides, useful for wound healing or tissue regeneration.
PT stimulating implant fixation and angiogenesis, and for treating and/or
PT preventing tumor.
XX
PS Disclosure; SEQ ID NO 7; 19pp; English.
XX
CC The invention provides criptin growth factor (CGF) polypeptides and their
CC encoding polynucleotides. The invention is useful for wound healing and
CC tissue regeneration, stimulating implant fixation, angiogenesis and for
CC treating and preventing neoplasia such as tumour. The invention is also
CC useful in gene therapy. The present sequence is human cripto growth
CC factor protein. This sequence is used in the invention.
XX
SQ Sequence 174 AA;
  Query Match      100.0%; Score 211; DB 8; Length 174;
  Best Local Similarity 100.0%; Pred. No. 2.5e-13;
  Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ELNRTCCCLNGGTCMLGSCFACACPPSPFYGRNCEHDVR 35
  |||||
DB 63 ELNRTCCCLNGGTCMLGSCFACACPPSPFYGRNCEHDVR 97

RESULT 8
ADS88699
ID ADS88699 standard; protein; 183 AA.
XX
AC ADS88699;
XX
```

---

```
DT 16-DEC-2004 (first entry)
XX
DE Amino acid sequence of a human secreted His-tagged Cripto protein.
XX
KW stem cell differentiation; cardiomyocyte; stem cell;
KW epidermal growth factor; EGF; EGF-CFC family; Cripto; neuronal cell;
KW Alq4; Nodal; heart disease; neuropathology; myocardial infarction;
KW Parkinson's disease; Alzheimer's disease; retinal degeneration;
KW cell therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 174..183
FT /note="His tag"
XX
PN WO2004083375-A2.
XX
PD 30-SEP-2004.
XX
PF 19-MAR-2004; 2004WO-IT000133.
XX
PR 21-MAR-2003; 2003IT-RM000125.
PR 29-JUL-2003; 2003IT-RM000370.
XX
PA (MINC/) MINCHIOTTI G.
PA (PERS/) PERSICO M.
PA (PARI/) PARISI S.
XX
PI Minchiotti G, Persico M, Parisi S;
XX
DR WPI; 2004-691035/67.
XX
PT Inducing stem cell differentiation in cardiomyocytes or neuronal cells by
PT exposure to a protein of EGF-CFC or Cripto protein inhibitors,
PT respectively, useful in treating heart diseases and neuropathologies.
XX
PS Disclosure; Page 13; 35pp; English.
XX
CC The specification describes a method for inducing stem cell
CC differentiation into cardiomyocytes. The method comprises exposing the
CC cells to a protein of the epidermal growth factor (EGF)-CFC family or its
CC derivatives and having at least the EGF and CFC domains (e.g. Cripto
CC protein). A second method is described for stem cell differentiation into
CC neuronal cells, where the cells are exposed to an inhibitor of the Cripto
CC protein. The Cripto protein inhibitor is an anti-Cripto antibody or its
CC functional fragments, or is a peptide specifically selected from a random
CC combinatorial peptide library, or is an antagonist of the Alq4(receptor)-
CC Cripto(co-receptor)-Nodal(ligand) pathway. The differentiated stem cells
CC of the invention are useful in the treatment of heart diseases and
CC neuropathologies, e.g. myocardial infarction, Parkinson's disease,
CC Alzheimer's disease and retinal degeneration. They are especially useful
CC for cell therapy. The Cripto protein or its inhibitors is useful in the
CC preparation of a composition able to direct stem cell differentiation
CC toward the neuronal lineage. The present sequence represents a Cripto
CC protein, which maybe used in the method of the invention.
XX
SQ Sequence 183 AA;
  Query Match      100.0%; Score 211; DB 8; Length 183;
  Best Local Similarity 100.0%; Pred. No. 2.6e-13;
  Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ELNRTCCCLNGGTCMLGSCFACACPPSPFYGRNCEHDVR 35
  |||||
DB 77 ELNRTCCCLNGGTCMLGSCFACACPPSPFYGRNCEHDVR 111

RESULT 9
AAR22548
ID AAR22548 standard; protein; 188 AA.
XX
```

AC AAR22548;  
 XX 25-MAR-2003 (revised)  
 DT 17-DEC-2001 (revised)  
 DT 28-AUG-1992 (first entry)  
 XX  
 DE Human CRIPTO CR-1 protein.  
 XX  
 XX cell proliferation; tumour; CR-3; transforming growth factor;  
 KW epidermal growth factor.  
 XX  
 OS Homo sapiens.  
 XX  
 PN USN7749001-N.  
 XX  
 PD 25-FEB-1992.  
 XX  
 PF 23-AUG-1991; 91US-00749001.  
 XX  
 PR 23-AUG-1991; 91US-00749001.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
 XX  
 XX Salomon D, Persico M;  
 PI  
 XX WPI; 1992-123675/15.  
 DR  
 DR N-PSDB; AAQ22495.  
 XX  
 XX New crypto gene CR-1 and crypto-related gene CR-3 genomic DNA - CR-3  
 PT protein and anti-CR-3 antibodies, useful in immunoassay to detect CR-3 as  
 PT tumour specific marker.  
 XX  
 PS Disclosure; Page 22; 44pp; English.  
 XX  
 XX The CR-1 genomic clone which codes for this protein can be used in  
 CC transgenic animals to examine the effects of overexpression of this gene  
 CC on development and tumorigenicity and to study regulation of CR-1 gene.  
 CC See also AAQ22494. (Note: Revised entry submitted to correct the patent  
 CC number format of US Government-owned NTIS applications to prevent clashes  
 CC with ongoing US granted patent numbers. For further information please  
 CC visit the Derwent web site at [www.derwent.com/dwpi/updates/ntis\\_us.html](http://www.derwent.com/dwpi/updates/ntis_us.html).)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 XX Sequence 188 AA;  
 SQ  
 Query Match 100.0%; Score 211; DB 2; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-13;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ELNRTCLLGGTCTMLGSPFCACPPSPFYGRNCEHDVR 35  
 DB 77 ELNRTCLLGGTCTMLGSPFCACPPSPFYGRNCEHDVR 111  
 RESULT 10  
 AAW29735  
 ID AAW29735 standard; protein; 188 AA.  
 AC AAW29735;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 09-NOV-1998 (first entry)  
 XX  
 XX Homo sapiens CRIPTO protein.  
 DE  
 XX CRIPTO; diagnosis; prognosis; cancer; colorectal; breast.  
 KW  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH Cleavage-site 28..29  
 FT Cleavage-site /note="V-A"  
 FT Region 79..81

FT Cleavage-site /note="potential Asp glycosylation sequence"  
 FT 111..112  
 FT /note="R-K"  
 FT 126..127  
 FT /note="K-K"  
 FT 159..160  
 FT /note="V-A"  
 FT 171..174  
 FT /note="R-T-T-T"  
 XX  
 PN US5792616-A.  
 XX  
 PD 11-AUG-1998.  
 XX  
 PF 05-JUN-1995; 95US-00463616.  
 XX  
 XX 29-MAY-1990; 90US-00530165.  
 PR 28-SEP-1992; 92US-00947315.  
 PR 10-NOV-1994; 94US-00337911.  
 XX  
 PA (USGO ) US GOVERNMENT.  
 XX  
 XX Persico MG, Salomon DS;  
 PI  
 XX WPI; 1998-456123/39.  
 DR  
 DR N-PSDB; AAV47553.  
 XX  
 XX New antibody which binds a CRIPTO protein - is useful for screening for  
 PT expression of a CRIPTO protein in a tissue sample.  
 PT  
 PS Disclosure; Fig 2; 12pp; English.  
 XX  
 XX The sequence is that of the human CRIPTO protein. Expression of CRIPTO  
 CC mRNA and CRIPTO protein appears to be associated with various cancers,  
 CC e.g. colon, breast and small cell lung carcinomas and thus the sequence  
 CC or antibodies raised against it may be of use in the diagnosis of  
 CC cancers. (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 XX Sequence 188 AA;  
 SQ  
 Query Match 100.0%; Score 211; DB 2; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-13;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ELNRTCLLGGTCTMLGSPFCACPPSPFYGRNCEHDVR 35  
 DB 77 ELNRTCLLGGTCTMLGSPFCACPPSPFYGRNCEHDVR 111  
 RESULT 11  
 AAW87630  
 ID AAW87630 standard; protein; 188 AA.  
 XX  
 AC AAW87630;  
 XX  
 XX 03-MAR-1999 (first entry)  
 DT  
 DE Human CRIPTO-related (CR-1) protein.  
 XX  
 XX human CRIPTO-related protein; CR-1; antibody; CR-3; proliferation;  
 KW differentiation; transformation; mesenchymal cell; epithelial cell;  
 KW tumour specific marker; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5854399-A.  
 XX  
 PD 29-DEC-1998.  
 XX  
 PF 05-JUN-1995; 95US-00464023.  
 XX  
 XX 23-AUG-1991; 91US-00749001.  
 PR 17-NOV-1993; 93US-00154198.

```

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Persica MG, Salomon DS;
XX
XX WPI; 1999-095001/08.
XX N-PSDB; AAV83923.
XX
XX New antibody to human CRIPTO-related polypeptide-3 - used for the
XX detection of CRIPTO-related polypeptide-3 which is involved in the
XX regulation of the proliferation, differentiation and transformation of
XX cells.
XX
XX Example 2; Col 17-18; 26pp; English.
XX
XX The present sequence represents a human CRIPTO-related protein CR-1. The
XX specification describes an antibody which has binding affinity to CR-3
XX and not to CR-1. The antibodies can be used for the detection of CR-3. CR
XX -3 is a regulatory molecule involved in regulating the proliferation,
XX differentiation, and transfection of various mesenchymal and epithelial
XX cells. In addition, expression of CR-3 may serve as a tumour specific
XX marker that may have applicability in the diagnosis, prognosis and
XX possible treatment of specific types of cancer
XX
XX Sequence 188 AA;
XX
Query Match 100.0%; Score 211; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 2.7e-13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ELNRTCCLLGGTCTMLGSCFACPPSPFYGRNCEHDVR 35
DB 77 ELNRTCCLLGGTCTMLGSCFACPPSPFYGRNCEHDVR 111
XX
RESULT 12
ABB77101
ID ABB77101 standard; protein; 188 AA.
XX
AC ABB77101;
XX
DT 08-OCT-2002 (first entry)
XX
DE Human CripTo-1 full length protein.
XX
KW Human; CripTo-1; CR-1; mutant; tumour; cell proliferation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..30
FT /label= signal_peptide
FT Protein 31..188
FT /label= Mature_CripTo-1
XX
PN WO200222808-A2.
XX
PD 21-MAR-2002.
XX
PF 18-SEP-2001; 2001WO-US029066.
XX
PR 18-SEP-2000; 2000US-0233148P.
XX
PA (BIOU ) BIOGEN INC.
XX
PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;
XX
DR WPI; 2002-339868/37.
XX N-PSDB; ABL55852.
XX
XX New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful
XX for treating cell proliferation, especially cancer, comprises amino acid
XX PT change that prevents fucosylation at Thr88.
XX

```

---

```

XX Claim 2; Fig 1; 41pp; English.
XX
XX The sequence represents the full length human CripTo-1 protein. The
XX invention relates to a novel mutant CripTo polypeptide, or its functional
XX fragment, having at least one amino acid alteration at positions 86, 87
XX or 88. The mutant polypeptide, or its chimera, is used to inhibit growth
XX of tumour cells, in vivo or in vitro, particularly for treating breast,
XX ovarian, renal, colorectal, uterine, prostatic, lung, bladder or central
XX nervous system cancers, melanoma and leukaemia, also generally for
XX treating undesired cell proliferation
XX
XX Sequence 188 AA;
XX
Query Match 100.0%; Score 211; DB 5; Length 188;
Best Local Similarity 100.0%; Pred. No. 2.7e-13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ELNRTCCLLGGTCTMLGSCFACPPSPFYGRNCEHDVR 35
DB 77 ELNRTCCLLGGTCTMLGSCFACPPSPFYGRNCEHDVR 111
XX
RESULT 13
AAO14638
ID AAO14638 standard; protein; 188 AA.
XX
AC AAO14638;
XX
DT 13-JUN-2002 (first entry)
XX
DE Human cripTo protein.
XX
KW Human; cripTo protein; epidermal growth factor family; EGF family;
XX gene therapy; protein therapy; vaccine; lung cancer; colon cancer;
XX colorectal cancer; breast cancer.
XX
OS Homo sapiens.
XX
PN WO200216413-A2.
XX
PD 28-FEB-2002.
XX
PF 20-AUG-2001; 2001WO-EP009646.
XX
PR 24-AUG-2000; 2000GB-00020953.
XX
PR (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cassart J, Coche T, Palmantier RM, Vinals Y De BassolsC;
XX
XX WPI; 2002-280910/32.
XX DR N-PSDB; AAL42198.
XX
XX New CripTo tumor polynucleotides, useful in medicine, particularly for
XX treating tumors (e.g. lung, colorectal, colon or breast tumor) that
XX express a cripTo antigen.
XX
XX Disclosure; Page 95; 134pp; English.
XX
XX The invention comprises the nucleotide and amino acid sequences of human
XX cripTo proteins. CripTo is a 188 amino acid protein that shares
XX homologies with the epidermal growth factor (EGF) family. The cripTo
XX nucleic acid and protein sequences of the invention are useful for
XX treating (e.g. gene therapy, protein therapy, vaccine) cancer, such as:
XX lung, colon, colorectal or breast. The present amino acid sequence
XX represents a human cripTo protein
XX
XX Sequence 188 AA;
XX
Query Match 100.0%; Score 211; DB 5; Length 188;
Best Local Similarity 100.0%; Pred. No. 2.7e-13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX

```

Qy 1 ELNRTCLNGGTCMLGSPFCACPPSPFYGRNCEHDVR 35  
Db 77 ELNRTCLNGGTCMLGSPFCACPPSPFYGRNCEHDVR 111

RESULT 14  
AAO14636  
ID AAO14636 standard; protein; 188 AA.

XX AAO14636;  
XX 13-JUN-2002 (first entry)  
XX Human cripto 1 protein.

XX Human; cripto 1 protein; epidermal growth factor family; EGF family;  
KW gene therapy; protein therapy; vaccine; lung cancer; colon cancer;  
KW colorectal cancer; breast cancer.

XX Homo sapiens.  
XX WO200216413-A2.

XX 28-FEB-2002.

XX 20-AUG-2001; 2001WO-EP009646.

XX 24-AUG-2000; 2000GB-00020953.

XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX Cassart J, Coche T, Palmantier RM, Vinals Y De BassolsC;

DR WPI; 2002-280910/32.

DR N-PSDB; AAL42196.

XX New Cripto tumor polynucleotides, useful in medicine, particularly for  
PT treating tumors (e.g. lung, colorectal, colon or breast tumor) that  
PT express a cripto antigen.

XX Claim 1; Page 94; 134pp; English.

XX The invention comprises the nucleotide and amino acid sequences of human  
CC cripto proteins. Cripto is a 188 amino acid protein that shares  
CC homologies with the epidermal growth factor (EGF) family. The cripto  
CC nucleic acid and protein sequences of the invention are useful for  
CC treating (e.g. gene therapy, protein therapy, vaccine) cancer, such as:  
CC lung, colon, colorectal or breast. The present amino acid sequence  
CC represents the human cripto 1 protein

XX Sequence 188 AA;

Query Match 100.0%; Score 211; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 2.7e-13;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELNRTCLNGGTCMLGSPFCACPPSPFYGRNCEHDVR 35  
Db 77 ELNRTCLNGGTCMLGSPFCACPPSPFYGRNCEHDVR 111

RESULT 15  
AAO14727  
ID AAO14727 standard; protein; 188 AA.

XX AAO14727;

XX 13-JUN-2002 (first entry)

XX Human variant cripto 1 protein.

XX Human; variant cripto 1 protein; epidermal growth factor family;

KW EGF family; gene therapy; protein therapy; vaccine; lung cancer;  
KW colon cancer; colorectal cancer; breast cancer; mutant; mutein.  
XX Homo sapiens.  
OS Synthetic.

XX Key Location/Qualifiers  
FT Misc-difference 22 /note= "wild type val replaced by Ala"  
FT

XX WO200216413-A2.

XX 28-FEB-2002.

XX 20-AUG-2001; 2001WO-EP009646.

XX 24-AUG-2000; 2000GB-00020953.

XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX Cassart J, Coche T, Palmantier RM, Vinals Y De BassolsC;

XX WPI; 2002-280910/32.

XX N-PSDB; AAL42204.

XX New Cripto tumor polynucleotides, useful in medicine, particularly for  
PT treating tumors (e.g. lung, colorectal, colon or breast tumor) that  
PT express a cripto antigen.

XX Example 2; Page 105; 134pp; English.

XX The invention comprises the nucleotide and amino acid sequences of human  
CC cripto proteins. Cripto is a 188 amino acid protein that shares  
CC homologies with the epidermal growth factor (EGF) family. The cripto  
CC nucleic acid and protein sequences of the invention are useful for  
CC treating (e.g. gene therapy, protein therapy, vaccine) cancer, such as:  
CC lung, colon, colorectal or breast. The present amino acid sequence  
CC represents a variant human cripto 1 protein

XX Sequence 188 AA;

Query Match 100.0%; Score 211; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 2.7e-13;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELNRTCLNGGTCMLGSPFCACPPSPFYGRNCEHDVR 35

Db 77 ELNRTCLNGGTCMLGSPFCACPPSPFYGRNCEHDVR 111

Search completed: February 4, 2006, 14:32:25  
Job time : 141 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 14:32:47 ; Search time 25 Seconds  
(without alignments)  
134.703 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_77\_111  
Perfect score: 211  
Sequence: 1 ELNRTCLNGGTCMLGSCACPPSYGRNCEHDVR 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211	100.0	188	2 A30362	teratocarcinoma-de
2	203	96.2	188	2 A39787	teratocarcinoma-de
3	159	94.3	171	2 I49612	teratocarcinoma-de
4	105	49.8	379	2 A59180	wnt inhibitory fac
5	97	46.0	4351	2 T00252	MEGF1 protein - ra
6	96.5	45.7	2524	2 A35844	Xotch protein - Af
7	96	45.5	378	2 B59180	wnt inhibitory fac
8	94.5	44.8	861	2 A48825	Notch homolog Motc
9	94.5	44.8	2331	2 S18188	notch protein homo
10	94.5	44.8	2531	2 A46019	notch-1 protein -
11	93.5	44.3	728	2 I50719	C-Delta-1 - chicke
12	92.5	43.8	1429	2 S06434	homotetic protein l
13	92.5	43.8	2437	2 S42612	transmembrane prot
14	92	43.6	835	2 JP0076	nel protein - chic
15	91.5	43.4	2555	2 A40043	notch protein homo
16	90.5	42.9	655	1 A46688	hepatocyte growth
17	90.5	42.9	685	2 JC7570	Delta-4 protein -
18	90.5	42.9	2703	1 A24420	Notch protein - fr
19	89.5	42.4	387	2 B49175	Motch A protein -
20	88.5	41.9	686	2 JC7569	Delta-4 protein -
21	88.5	41.9	2352	2 T30201	Notch homolog prot
22	88.5	41.9	3562	2 A47171	chondroitin sulfat
23	88	41.7	433	1 UKBAY	u-plasminogen acti
24	87	41.2	1203	2 A49175	Motch B protein -
25	87	41.2	2471	2 A49128	cell-fate determin
26	86.5	41.0	293	2 B26637	neurogenic repetit
27	86.5	41.0	1364	2 T03059	notch4 - mouse
28	86.5	41.0	2139	2 A35672	crumbs protein - f
29	86.5	41.0	3871	2 T22812	hypothetical prote

RESULT 1

A30362  
teratocarcinoma-derived growth factor 1 - human  
N:Alternate names: CRIPTO protein  
C:Species: Homo sapiens (man)  
C:Date: 18-Oct-1989 #sequence\_revision 18-Oct-1989 #text\_change 09-Jul-2004  
C:Accession: B39787; A30362  
R:Dono, R.; Montuori, N.; Rocchi, M.; De Ponti-Zilli, L.; Ciccodicola, A.; Persico, M.G.  
Am. J. Hum. Genet. 49, 555-565, 1991  
A:Title: Isolation and characterization of the CRIPTO autosomal gene and its X-linked re  
A:Reference number: A39787; MUID:91353571; PMID:1882841  
A:Accession: B39787  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-188 <DON>  
A:Cross-references: UNIPROT:P13385; UNIPARC:UPI000004966D; GB:M96955; GB:M37099; NID:G33;  
R:Ciccodicola, A.; Dono, R.; Obici, S.; Simeone, A.; Zollo, M.; Persico, M.G.  
EMBO J. 8, 1987-1991, 1989  
A:Title: Molecular characterization of a gene of the 'EGF family' expressed in undifferer  
A:Reference number: A30362; MUID:90005403; PMID:2792079  
A:Accession: A30362

ALIGNMENTS

30	85.5	40.5	559	1	A29941	t-plasminogen acti
31	85.5	40.5	559	1	A35029	t-plasminogen acti
32	85.5	40.5	1064	2	A40136	fibropellin Ia - s
33	85.5	40.5	1125	1	S57846	protein-tyrosine k
34	85.5	40.5	2321	2	S78549	notch3 protein - h
35	85	40.3	2871	2	A55567	fibrellin I - bovi
36	85	40.3	2871	2	A55624	fibrellin-1 precur
37	85	40.3	3002	2	A47221	fibrellin 1 precur
38	84.5	40.0	461	1	KFHU	coagulation factor
39	84.5	40.0	2318	2	S45306	notch 3 protein -
40	84	39.8	431	1	UKHU	u-plasminogen acti
41	84	39.8	1251	2	A57293	latent transformin
42	84	39.8	1574	2	T13954	MEGF6 protein - ra
43	83.5	39.6	407	1	KF807	coagulation factor
44	83	39.3	427	2	JC4915	ags protein precur
45	83	39.3	810	2	T10756	Nel-homolog protei

Query Match 100.0%; Score 211; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.2e-17;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCACPPSYGRNCEHDVR 35  
|||||  
DB 77 ELNRTCLNGGTCMLGSCACPPSYGRNCEHDVR 111

RESULT 2

A39787  
teratocarcinoma-derived growth factor - human  
C:Species: Homo sapiens (man)  
C:Date: 20-Mar-1992 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
R:Dono, R.; Montuori, N.; Rocchi, M.; De Ponti-Zilli, L.; Ciccodicola, A.; Persico, M.G.  
Am. J. Hum. Genet. 49, 555-565, 1991  
A:Title: Isolation and characterization of the CRIPTO autosomal gene and its X-linked re  
A:Reference number: A39787; MUID:91353571; PMID:1882841  
A:Accession: A39787  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-188 <DON>  
A:Cross-references: UNIPROT:Q8TCC1; UNIPARC:UPI0000145185

A;Note: the authors translated the codon GAC for residue 43 as Glu  
C;Superfamily: teratocarcinoma-derived growth factor 1; EGF homology

Query Match 96.2%; Score 203; DB 2; Length 188;  
Best Local Similarity 97.1%; Pred. No. 1e-16;  
Matches 34; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACPPSPFYGRNCEHDVR 35  
|||||  
DB 77 ELNRTCLNGGTCMLGSCFACPPSPFYGRNCEHDVR 111  
|||||

RESULT 3  
I49612  
teratocarcinoma-derived growth factor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I49612  
R;Dono, R.; Scalaria, L.; Pacifico, F.; Simeone, A.; Persico, M.G.; Acampora, D.  
Development 118, 1157-1168, 1993  
A;Title: The murine criptor gene: expression during mesoderm induction and early heart  
A;Reference number: I49612; MUID:94034736; PMID:7916676  
A;Accession: I49612  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-171 <RES>  
A;Cross-references: UNIPROT:P51865; UNIPARC:UPI0000029326; GB:M87321; NID:G402714; PID:  
C;Genetics:  
A;Gene: cripto  
C;Superfamily: teratocarcinoma-derived growth factor 1; EGF homology

Query Match 94.3%; Score 199; DB 2; Length 171;  
Best Local Similarity 94.1%; Pred. No. 2.8e-16;  
Matches 32; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNRTCLNGGTCMLGSCFACPPSPFYGRNCEHDVR 35  
|||||  
DB 62 LNRTCLNGGTCMLGSCFACPPSPFYGRNCEHDVR 95  
|||||

RESULT 4  
A59180  
Wnt inhibitory factor-1 - human  
C;Species: Homo sapiens (man)  
C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 09-Jul-2004  
C;Accession: A59180  
R;Hsieh, J.C.; Kodjabachian, L.; Rebbert, M.L.; Rattner, A.; Smallwood, P.M.; Samos, C.  
Nature 398, 431-436, 1999  
A;Title: A new secreted protein that binds to Wnt proteins and inhibits their activities  
A;Reference number: A59180; MUID:99215557; PMID:10201374  
A;Accession: A59180  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-379 <HSI>  
A;Cross-references: UNIPROT:Q9Y5W5; UNIPARC:UPI0000051058; GB:AF122922; NID:G4585369; PTD:

Query Match 49.8%; Score 105; DB 2; Length 379;  
Best Local Similarity 60.0%; Pred. No. 4.6e-05;  
Matches 15; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 7 CLNGGTCMLGSCFACPPSPFYGRNCE 31  
|||||  
DB 218 CMNGGLCVTPGFCICPPGFYGVNCD 242  
|||||

RESULT 5  
T00252  
MEGF1 protein - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 09-Jul-2004  
C;Accession: T00252  
R;Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
Genomics 51, 27-34, 1998

A;Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
A;Reference number: Z14126; MUID:98360089; PMID:9693030

A;Accession: T00252

A;Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A;Residues: 1-4351 <NAK>

A;Cross-references: UNIPROT:O98277; UNIPARC:UPI0000055B23; EMBL:AB011527; NID:g3449285; I

A;Experimental source: brain

C;Genetics:

A;Gene: MEGF1

C;Superfamily: rat MEGF1 protein; cadherin repeat homology; EGF homology; laminin G repeat

F;3798-3949/Domain: laminin G repeat homology <LGR>

F;3953-3985/Domain: EGF homology <EGF>

F;3992-4023/Domain: EGF homology <EGF1>

Query Match 46.0%; Score 97; DB 2; Length 4351;  
Best Local Similarity 65.5%; Pred. No. 0.0025;  
Matches 19; Conservative 1; Mismatches 5; Indels 4; Gaps 2;

QY 7 CLNGGTC--WLGS--FCACPPSFVGNRCE 31  
|||||: |||| | |||| | ||||  
Db 3958 CLNGGCSPALGGYLRCPCPPPSGRNCE 3986

RESULT 6

A35844

Xotch protein - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 12-Oct-1990 #sequence\_revision 12-Oct-1990 #text\_change 05-Oct-2004

C;Accession: A35844

R;Coffman, C.; Harris, W.; Kintner, C.  
Science 249, 1438-1441, 1990

A;Title: Xotch, the Xenopus homolog of Drosophila notch.

A;Reference number: A35844; MUID:90385285; PMID:2402639

A;Accession: A35844

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translat

A:Molecule type: mRNA

A;Residues: 1-2524 <COF>

A;Cross-references: UNIPARC:UPI000004P253

C;Superfamily: notch protein; ankryrin repeat homology; EGF homology

C;Keywords: transmembrane protein

F;146-177/Domain: EGF homology <EGX1>

F;184-215/Domain: EGF homology <EGF1>

F;222-254/Domain: EGF homology <EGF>

F;456-487/Domain: EGF homology <EGX2>

F;757-788/Domain: EGF homology <EGF3>

F;1025-1056/Domain: EGF homology <EGX3>

F;1924-1956/Domain: ankryrin repeat homology <AN1>

F;1957-1989/Domain: ankryrin repeat homology <AN2>

F;1991-2023/Domain: ankryrin repeat homology <AN3>

F;2024-2056/Domain: ankryrin repeat homology <AN4>

F;2057-2089/Domain: ankryrin repeat homology <AN5>

Query Match 45.7%; Score 96.5; DB 2; Length 2524;  
Best Local Similarity 43.2%; Pred. No. 0.0019;  
Matches 16; Conservative 7; Mismatches 11; Indels 3; Gaps 2;

QY 1 ELNRCTCLNGTGM--LGSP-CACPPSFGVRNCHEHDV 34  
: : |||||: : |||||: |||||: ||:  
Db 986 DTESSCFNAGTGCDGINFTCCQPFPPTGSGCHDI 1022

RESULT 7

B59180

Wnt inhibitory factor-1 - zebra fish

C;Species: Brachydanio rerio (zebra fish)

C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 09-Jul-2004

C;Accession: B59180

R;Haieh, J.C.; Kodjabachian, L.; Rabbert, M.L.; Rattner, A.; Smallwood, P.M.; Samos, C.H.  
Nature 398, 431-436, 1999

A;Title: A new secreted protein that binds to Wnt proteins and inhibits their activities.

A;Reference number: A59180; MUID:99215557; PMID:10201374

A;Accession: B59180

A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-378 <HSA>  
A:Cross-references: UNIPROT:Q9W6F9; UNIPARC:UPI0000138F08; GB:AF122925; NID:g4585375; PID:122925; UNIPROT:Q9W6F9; UNIPARC:UPI0000138F08; GB:AF122925; NID:g4585375; PID:122925

Query Match 45.5%; Score 96; DB 2; Length 378;  
Best Local Similarity 56.0%; Pred. No. 0.00051;  
Matches 14; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 7 CLNGGTCMLGSCACPPSPFYGRNCE 31  
Db 216 CLNGGLCMSPGVVICPPGYFGSSCE 240

RESULT 8  
A48825  
Notch homolog Notch protein - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 01-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Oct-2004  
C:Accession: A48825  
R:Reaume, A.G.; Conlon, R.A.; Zirngibl, R.; Yamaguchi, T.P.; Roseant, J.  
Dev. Biol. 154; 377-387, 1992  
A:Title: Expression analysis of a Notch homologue in the mouse embryo.  
A:Reference number: A48825; MUID:93050801; PMID:11426644  
A:Accession: A48825  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-861 <REA>  
A:Cross-references: UNIPARC:UPI000017745F  
A:Experimental source: embryo  
A:Note: sequence extracted from NCBI backbone (NCBIP:119144)  
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology  
F:26-57/Domain: EGF homology <EGP>  
F:64-95/Domain: EGF homology <EGX1>  
F:198-229/Domain: EGF homology <EGF2>  
F:441-472/Domain: EGF homology <EGX2>

Query Match 44.8%; Score 94.5; DB 2; Length 861;  
Best Local Similarity 45.9%; Pred. No. 0.0014;  
Matches 17; Conservative 6; Mismatches 11; Indels 3; Gaps 2;

Qy 1 ELNRTCLLNGGTGM--LGSF-CACPPSPFYGRNCEHDV 34  
Db 197 DCTESSCFNGGTCVDGINSFTCLCPGFTGTCYQYDV 233

RESULT 9  
S18188  
notch protein homolog - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 02-Aug-2002  
C:Accession: S18188  
R:Weinmaster, G.; Roberts, V.J.; Lemke, G.  
Development 113; 199-205, 1991  
A:Title: A homolog of Drosophila Notch expressed during mammalian development.  
A:Reference number: S18188; MUID:92111383; PMID:1764995  
A:Accession: S18188  
A:Molecule type: mRNA  
A:Residues: 1-2531 <WEI>  
A:Cross-references: UNIPARC:UPI0000177456; EMBL:X57405; NID:g57634; PID:g57635  
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology  
F:987-1018/Domain: EGF homology <EGF1>  
F:1025-1056/Domain: EGF homology <EGF2>  
F:1233-1264/Domain: EGF homology <EGF2>  
F:1917-1949/Domain: ankyrin repeat homology <ANI>  
F:1950-1982/Domain: ankyrin repeat homology <AN2>  
F:1984-2016/Domain: ankyrin repeat homology <AN3>  
F:2017-2049/Domain: ankyrin repeat homology <AN4>  
F:2050-2082/Domain: ankyrin repeat homology <AN5>

Query Match 44.8%; Score 94.5; DB 2; Length 2531;

F:757-788/Domain: EGF homology <EG09>  
F:795-826/Domain: EGF homology <EG10>  
F:873-904/Domain: EGF homology <EG11>  
F:911-942/Domain: EGF homology <EG12>  
F:949-980/Domain: EGF homology <EG13>  
F:987-1018/Domain: EGF homology <EG14>  
F:1025-1056/Domain: EGF homology <EG15>  
F:1063-1094/Domain: EGF homology <EG16>  
F:1149-1180/Domain: EGF homology <EG17>  
F:1187-1218/Domain: EGF homology <EG18>  
F:1233-1264/Domain: EGF homology <EGF4>  
F:1352-1383/Domain: EGF homology <EGF9>  
F:1391-1425/Domain: EGF homology <EGF>  
F:1917-1948/Domain: ankyrin repeat homology <AN1>  
F:1949-1981/Domain: ankyrin repeat homology <AN2>  
F:1983-2015/Domain: ankyrin repeat homology <AN3>  
F:2016-2048/Domain: ankyrin repeat homology <AN4>  
F:2049-2081/Domain: ankyrin repeat homology <AN5>

Query Match 44.8%; Score 94.5; DB 2; Length 2531;  
Best Local Similarity 45.9%; Pred. No. 0.0033;  
Matches 17; Conservative 6; Mismatches 11; Indels 3; Gaps 2;

QY 1 ELNRTCCCLNGGTCM--LGSPF-CACPPSFYGRNCEHDV 34  
DB 986 DCTESSCFNGGTCVDGINSFTCLCPFGTGSYCQYDV 1022

RESULT 11  
IS0719  
C-Delta-1 - chicken  
C:Species: Gallus gallus (chicken)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C:Accession: IS0719  
R:Henique, D.; Adam, J.; Myat, A.; Chitnie, A.; Lewis, J.; Ish-Horowitz, D.  
Nature 375, 787-790, 1995  
A>Title: Expression of a Delta homologue in prospective neurons in the chick.  
A:Reference number: IS0719; MUID:195319507; PMID:7596411  
A:Accession: IS0719  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-728 <HEN>  
A:Cross-references: UNIPROT:Q90656; UNIPARC:UPI00000FD1E5; EMBL:U26590; NID:g882411; PID  
C:Superfamily: delta-4 protein; EGF homology  
F:299-332/Domain: EGF homology <EGX1>  
F:339-370/Domain: EGF homology <EGF1>  
F:416-447/Domain: EGF homology <EGF2>  
F:454-485/Domain: EGF homology <EGF>  
F:492-523/Domain: EGF homology <EGF3>

Query Match 44.3%; Score 93.5; DB 2; Length 728;  
Best Local Similarity 50.0%; Pred. No. 0.0016;  
Matches 17; Conservative 5; Mismatches 9; Indels 3; Gaps 2;

QY 1 ELNRTCCCLNGGTC--MLGSPF-CACPPSFYGRNCE 31  
DB 338 ECDANPCNKGSGCTDLENSYCTCTCPGFGKNC 371

RESULT 12  
S06434  
homeotic protein lin-12 precursor - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 31-Dec-2004  
C:Accession: S06434; A24769  
R:Yochem, J.; Weston, K.; Greenwald, I.  
Nature 335, 547-550, 1988  
A>Title: The Caenorhabditis elegans lin-12 gene encodes a transmembrane protein with over  
A:Reference number: S06434; MUID:188334747; PMID:3419531  
A:Accession: S06434  
A:Molecule type: DNA  
A:Residues: 1-1429 <YOC>  
A:Cross-references: UNIPROT:P14585; UNIPARC:UPI00000610FE; EMBL:M12069; NID:g156357; PID

R:Greenwald, I.  
Cell 43, 583-590, 1985  
A:Reference number: A24769; MUID:86079540; PMID:3000611  
A:Accession: A24769  
A:Molecule type: DNA  
A:Residues: 173-712 <GRE>  
A:Cross-references: UNIPARC:UPI000017A0C4  
C:Genetics:  
A:Introns: 50/2; 90/1; 109/1; 172/3; 545/1; 589/2; 632/2; 1273/3; 1389/3  
C:Keywords: glycoprotein; transmembrane protein  
F:254-284/Domain: EGF homology <EGF1>  
F:507-540/Domain: EGF homology <EGF>  
F:547-578/Domain: EGF homology <EGF2>  
F:909-931/Domain: transmembrane #status predicted <TMM>  
F:1093-1125/Domain: ankyrin repeat homology <AN1>  
F:1206-1238/Domain: ankyrin repeat homology <AN2>  
F:1240-1272/Domain: ankyrin repeat homology <AN3>

Query Match 43.8%; Score 92.5; DB 2; Length 1429;  
Best Local Similarity 60.0%; Pred. No. 0.0036;  
Matches 15; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 7 CLNGGTCMLGSPF-CACPPSFYGRNCE 31  
DB 469 CFNGGKCLSG-FCVCPDPFTGNQCE 492

RESULT 13  
S42612  
transmembrane protein precursor - zebra fish  
C:Species: Brachydanio rerio (zebra fish)  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S42612  
R:Bierkamp, C.; Campos-Ortega, J.A.  
Mech. Dev. 43, 87-100, 1993  
A>Title: A zebrafish homologue of the Drosophila neurogenic gene Notch and its pattern of  
A:Reference number: S42612; MUID:94128602; PMID:8297791  
A:Accession: S42612  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2437 <BIE>  
A:Cross-references: UNIPROT:P46530; UNIPARC:UPI00001305E8; EMBL:X69088; NID:g433866; PID  
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology  
F:755-786/Domain: EGF homology <EGF1>  
F:1023-1054/Domain: EGF homology <EGF2>  
F:1185-1216/Domain: EGF homology <EGF>  
F:1915-1947/Domain: ankyrin repeat homology <AN1>  
F:1948-1980/Domain: ankyrin repeat homology <AN2>  
F:1982-2014/Domain: ankyrin repeat homology <AN3>  
F:2015-2047/Domain: ankyrin repeat homology <AN4>  
F:2048-2080/Domain: ankyrin repeat homology <AN5>

Query Match 43.8%; Score 92.5; DB 2; Length 2437;  
Best Local Similarity 45.9%; Pred. No. 0.0054;  
Matches 17; Conservative 5; Mismatches 12; Indels 3; Gaps 2;

QY 1 ELNRTCCCLNGGTCM--LGSPF-CACPPSFYGRNCEHDV 34  
DB 984 DCTESSCFNGGTCVDGINSFSCVCLPFGFTGNYCQHDV 1020

RESULT 14  
JP0076  
nel protein - chicken  
C:Species: Gallus gallus (chicken)  
C>Date: 24-Feb-1995 #sequence\_revision 24-Feb-1995 #text\_change 04-Apr-2004  
C:Accession: A38963; JP0076  
R:Matsumashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.; Taniguchi, S.; Hori, K.  
Dev. Dyn. 203, 212-222, 1995  
A>Title: New gene, nel, encoding a Mr 93K protein with EGF-like repeats is strongly expre  
A:Reference number: A38963; MUID:95383734; PMID:7655083  
A:Accession: A38963  
A:Molecule type: mRNA

A:Residues: 1-835 <MAT>  
 A:Cross-references: UNIPARC:UPI000017C023; DDBJ:D45365  
 A:Experimental source: 9-day embryo  
 R:Matsumashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.; Taniguchi, S.; Hori, K.  
 submitted to JIPID, January 1995  
 A:Description: A new gene, nel, encoding a Mr 93K protein with EGF-like repeats is stron  
 A:Reference number: JP0076  
 A:Accession: JP0076  
 A:Molecule type: mRNA  
 A:Residues: 1-835 <MA2>  
 A:Cross-references: UNIPARC:UPI000017C023; DDBJ:D45365  
 A:Experimental source: 9-day embryo  
 F:273-333/Domain: von Willebrand factor type C repeat homology <VWC>  
 F:395-592/Region: EGF-like repeats  
 F:444-480/Domain: EGF homology <EGF1>  
 F:486-521/Domain: EGF homology <EGF>  
 F:525-552/Domain: EGF homology <EGF2>

Query Match 43.6%; Score 92; DB 2; Length 835;  
 Best Local Similarity 50.0%; Pred. No. 0.0027;  
 Matches 14; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 7 CLNGGTCLGSGFCACPPSPFYGRNCEHDV 34  
 DB 529 CRNGGACIASNVCACPQGTGFCETDI 556

RESULT 15  
 A40043  
 notch protein homolog TAN-1 precursor - human  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Apr-1992 #sequence\_revision 21-Apr-1992 #text\_change 05-Oct-2004  
 C:Accession: A40043  
 R:Ellisen, L.W.; Bird, J.; West, D.C.; Soreng, A.L.; Reynolds, T.C.; Smith, S.D.; Sklar,  
 Cell 66, 649-661, 1991  
 A:Title: TAN-1, the human homolog of the Drosophila Notch gene, is broken by chromosomal  
 A:Reference number: A40043; MUID:91347367; PMID:1831892  
 A:Accession: A40043  
 A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
 A:Molecule type: mRNA  
 A:Residues: 1-2555 <ELL>  
 A:Cross-references: UNIPARC:UPI0000177455; GB:M73980  
 C:Superfamily: notch protein; ankyrin repeat homology; EGF homology  
 F:261-292/Domain: EGF homology <EGX1>  
 F:494-525/Domain: EGF homology <EGF1>  
 F:987-1018/Domain: EGF homology <EGX2>  
 F:1149-1180/Domain: EGF homology <EGF>  
 F:1187-1218/Domain: EGF homology <EGF3>  
 F:1233-1264/Domain: EGF homology <EGX3>  
 F:1927-1959/Domain: ankyrin repeat homology <AN1>  
 F:1960-1992/Domain: ankyrin repeat homology <AN2>  
 F:1994-2026/Domain: ankyrin repeat homology <AN3>  
 F:2027-2059/Domain: ankyrin repeat homology <AN4>  
 F:2060-2092/Domain: ankyrin repeat homology <AN5>

Query Match 43.4%; Score 91.5; DB 2; Length 2555;  
 Best Local Similarity 45.9%; Pred. No. 0.0073;  
 Matches 17; Conservative 5; Mismatches 12; Indels 3; Gaps 2;

QY 1 ELNRTCLNGTGM--LGSP-CACPPSPFYGRNCEHDV 34  
 DB 986 DCTESSCFNGTGVGINSFTCLCPGFTGSCQHV 1022

Search completed: February 4, 2006, 14:38:30  
 Job time : 26 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 14:27:51 ; Search time 152.5 Seconds  
(without alignments)  
161.925 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_77\_111

Perfect score: 211

Sequence: 1 ELNRTCLNGTGMGLSPFCACPPSPYGRNCBHDVR 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.\*

1: uniprot\_eprot.\*

2: uniprot\_trenbl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	211	100.0	188	1	TDGF1_HUMAN	P13385 homo sapien
2	211	100.0	188	2	Q8TCC1_HUMAN	Q8TCC1 homo sapien
3	205	97.2	181	2	Q58D57_BOVIN	Q58D57 bos taurus
4	203	96.2	188	1	TDGF2_HUMAN	P51864 homo sapien
5	199	94.3	171	1	TDGF1_MOUSE	P51865 mus musculus
6	199	94.3	171	2	Q7TQ06_MOUSE	Q7TQ06 mus musculus
7	185	87.7	66	2	Q91AT2_BRARE	Q91AT2 brachydanio
8	147	69.7	183	2	Q57517_BRARE	O57517 brachydanio
9	147	69.7	183	2	Q50415_BRARE	O50415 brachydanio
10	147	69.7	190	2	O57516_BRARE	O57516 brachydanio
11	145	68.7	190	2	Q800J2_XENLA	Q800J2 xenopus lae
12	145	68.7	190	2	Q91649_XENLA	Q91649 xenopus lae
13	138	65.4	193	2	Q918Q3_CHICK	Q918Q3 gallus gall
14	136	64.5	223	2	Q53T05_HUMAN	Q53T05 homo sapien
15	136	64.5	223	2	Q9GZR3_HUMAN	Q9GZR3 homo sapien
16	134	63.5	154	2	Q9JIB7_MOUSE	Q9JIB7 mus musculus
17	134	63.5	202	2	P97766_MOUSE	P97766 mus musculus
18	105	49.8	220	2	Q99KR2_MOUSE	Q99KR2 mus musculus
19	105	49.8	379	1	WIF1_HUMAN	Q9Y5W5 homo sapien
20	105	49.8	379	1	WIF1_MOUSE	Q9YUW1 mus musculus
21	105	49.8	379	1	WIF1_RAT	Q61N38 rattus norv
22	100.5	47.6	2067	2	Q59ED8_HUMAN	Q59ED8 homo sapien
23	100.5	47.6	2555	2	Q58XM3_HUMAN	Q58XM3 homo sapien
24	100.5	47.6	2853	2	Q25253_LUCCU	Q25253 lucilla cup
25	98.5	46.7	1551	2	Q9NGV4_DROME	Q9NGV4 drosophila
26	98.5	46.7	3396	2	Q9VM55_DROME	Q9VM55 drosophila
27	98	46.4	1654	2	Q69ZY7_MOUSE	Q69ZY7 mus musculus
28	98	46.4	4351	2	Q5P226_MOUSE	Q5P226 mus musculus
29	97.5	46.2	2656	2	Q9GNI3_PARLI	Q9GNI3 paracentrot
30	97	46.0	814	2	Q7XKL5_XENLA	Q7XKL5 xenopus lae
31	97	46.0	4351	1	FAT2_RAT	O88277 rattus norv

32	96.5	45.7	2524	1	NOTCH_XENLA	P21783 xenopus lae
33	96	45.5	354	2	QSPQ42_XENLA	QSPQ42 xenopus lae
34	96	45.5	374	1	WIF1_XENLA	Q9W6F8 xenopus lae
35	96	45.5	378	1	WIF1_BRARE	Q9W6F9 brachydanio
36	96	45.5	378	2	O5TZI0_BRARE	O5TZI0 brachydanio
37	96	45.5	664	1	DLLC_BRARE	Q9IAT6 brachydanio
38	95.5	45.3	963	2	Q9GPM9_CAEBR	Q9GPM9 caenorhabdi
39	95.5	45.3	1270	2	Q9GPM0_CAEBR	Q9GPM0 caenorhabdi
40	94.5	44.8	721	2	Q91902_XENLA	Q91902 xenopus lae
41	94.5	44.8	726	2	Q9AW87_CYNPY	Q9AW87 cynops pyrr
42	94.5	44.8	2516	2	O7TQ52_MOUSE	O7TQ52 mus musculus
43	94.5	44.8	2526	2	Q7TQ51_MOUSE	Q7TQ51 mus musculus
44	94.5	44.8	2531	1	NOTC1_MOUSE	Q01705 mus musculus
45	94.5	44.8	2531	1	NOTC1_RAT	Q07008 rattus norv

#### ALIGNMENTS

RESULT 1  
TDGF1\_HUMAN STANDARD; PRT; 188 AA.  
AC P13385;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Teratocarcinoma-derived growth factor 1 precursor (Epidermal growth factor-like crypto protein CRI) (Crypto-1 growth factor) (CRGF).  
GN Name=TDGF1; Synonyms=CRIPTO;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=90005403; PubMed=2792079;  
RA Ciccodicola A., Dono R., Obici S., Zollo M., Persico M.G.;  
RT "Molecular characterization of a gene of the 'EGF family' expressed in undifferentiated human NTera2 teratocarcinoma cells.";  
RL EMBO J. 8:1987-1991(1989).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=91353571; PubMed=1882841;  
RA Dono R., Montuori N., Rocchi M., de Ponti-Zilli L., Ciccodicola A., Persico M.G.;  
RT "Isolation and characterization of the CRIPTO autosomal gene and its X-linked related sequence.";  
RL Am. J. Hum. Genet. 49:555-565(1991).  
RN [3]  
RP PROTEIN SEQUENCE OF 31-45.  
RX PubMed=15340161; DOI=10.1110/ps.04682504;  
Zhang Z., Henzel W.J.;  
RT "Signal peptide prediction based on analysis of experimentally verified cleavage sites.";  
RL Protein Sci. 13:2819-2824(2004).  
RN [4]  
RX DISULFIDE BONDS.  
MEDLINE=22800780; PubMed=12919325;  
DOI=10.1046/j.1432-1033.2003.03749.x;  
Foley S.F., Van Vlijmen H.W., Boynton R.E., Adkins H.B., Cheung A.E., Singh J., Sanicola M., Young C.N., Wen D.;  
RT "The CRIPTO/FRL-1/CRYPRIC (CFC) domain of human Cripto.";  
RL Eur. J. Biochem. 270:3610-3618(2003).  
CC -!- FUNCTION: Could play a role in the determination of the epiblastic cells that subsequently give rise to the mesoderm.  
CC -!- TISSUE SPECIFICITY: Preferentially expressed in gastric and colorectal carcinomas than in their normal counterparts.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its





```

RN RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=21180013; PubMed=11282978; DOI=10.1101/gr.170101;
RA Smith T.P.L., Grosse W.M., Freking B.A., Roberts A.J., Stone R.T.,
RA Casas E., Wray J.E., White J., Cho J., Fahrenkrug S.C., Bennett G.L.,
RA Heaton M.P., Laegreid W.W., Rohrer G.A., Chitko-McKown C.G.,
RA Perle G., Holt I., Karamycheva S., Liang F., Quackenbush J.,
RA Keele J.W.;
RT "Sequence evaluation of four pooled-tissue normalized bovine cdna
RT libraries and construction of a gene index for cattle.";
RL Genome Res. 11:626-630(2001).
RN RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Pooled;
RA Harhay G.P., Sonstegard T.S., Van Tassel C.P., Clawson M.L.,
RA Heaton M.P., Keele J.W., Snelling W.M., Weidmann R.T., Smith T.P.L.;
RT "Sequencing and analysis of Bos taurus full-length insert cdna
RT clones.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BF021740; AAX46587.1; -; mRNA.
SQ SEQUENCE 181 AA; 20023 MW; A4E8AEA78E106514 CRC64;

Query Match 97.2%; Score 205; DB 2; Length 181;
Best Local Similarity 97.1%; Pred. No. 2.1e-17;
Matches 34; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACPPSPFYGRNCEHDVR 35
DB |||||
RESULT 4
TDGF2 HUMAN
ID TDGF2 HUMAN STANDARD; PRT; 188 AA.
AC P51864;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Teratocarcinoma-derived growth factor 2 (Epidermal growth factor-like
DE crypto protein CR3) (Crypto-3 growth factor).
GN Name=TDGF3; Synonyms=TDGF2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung fibroblast;
RX MEDLINE=91353571; PubMed=1882841;
RA Dono R., Montuori N., Rocchi M., de Ponti-Zilli L., Ciccodicola A.,
RA Persico M.G.;
RT "Isolation and characterization of the CRYPTO autosomal gene and its
RT X-linked related sequence.";
RL Am. J. Hum. Genet. 49:555-565(1991).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Johnson D., Wamsley P., Gibson A.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Kintner M.A., Kintner R.L., Hosick H.L.;
RT "Crypto-3 and Crypto-1 have different effects on the growth
RT characteristics of MCF-7 and Vero cells.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Could play a role in the determination of the epiblastic
CC cells that subsequently give rise to the mesoderm.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use as long as its content is in no way modified and this statement is not
CC removed.
DR EMBL; M96956; AAA61135.1; -; mRNA.
DR EMBL; AC000113; AAB46353.1; -; Genomic DNA.
DR EMBL; AF251549; AAG49538.1; -; Genomic DNA.
DR EMBL; AF251550; AAG49539.1; -; Genomic DNA.
DR HSSP; P00740; IEDM.
DR Ensembl; ENSG00000183514; Homo sapiens.
DR HGNC; HGNC:11703; TDGF3.
DR MIM; 187395; -.
DR GO; GO:0008083; P: growth factor activity; NAS.
DR GO; GO:0007500; P: mesodermal cell fate determination; NAS.
DR InterPro; IPR000742; EGF_2.
DR Pfam; PF00008; EGF; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00026; EGF_3; FALSE NEG.
KW EGF-like domain; Glycoprotein; Growth factor.
FT DOMAIN 78 107 EGF-like.
FT CARBOHYD 79 79 N-linked (GlcNAc...) (potential).
FT DISULFID 82 89 By similarity.
FT DISULFID 83 95 By similarity.
FT DISULFID 97 106 By similarity.
FT DISULFID 115 133 By similarity.
FT DISULFID 128 149 By similarity.
FT DISULFID 131 140 By similarity.
SQ SEQUENCE 188 AA; 21181 MW; C07AC973E4D82B32 CRC64;

Query Match 96.2%; Score 203; DB 1; Length 188;
Best Local Similarity 97.1%; Pred. No. 3.9e-17;
Matches 34; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACPPSPFYGRNCEHDVR 35
DB |||||
77 ELNRTCLNGGTCMLGSCFACPPSPFYGRNCEHDVR 111

RESULT 5
TDGF1 MOUSE
ID TDGF1 MOUSE STANDARD; PRT; 171 AA.
AC P51865;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Teratocarcinoma-derived growth factor precursor (Epidermal growth
DE factor-like Crypto protein) (Crypto growth factor).
GN Name=TDgf1; Synonyms=Cripto;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Embryo;
RX MEDLINE=94094736; PubMed=7916676;
RA Dono R., Scalera L., Pacifico F., Acampora D., Persico M.G.,
RA Simeone A.;
RT "The murine crypto gene: expression during mesoderm induction and
RT early heart morphogenesis.";
RL Development 118:1157-1168(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 135-171.
RC STRAIN=129/Sv;
RX MEDLINE=96269599; PubMed=8661720; DOI=10.1007/s003359800100;
RA Liguori G., Tucci M., Montuori N., Dono R., Lago C.T., Pacifico A.F.,
RA Persico M.G.;
RT "Characterization of the mouse Tdgf1 gene and Tdgf pseudogenes.";
RL Mamm. Genome 7:344-348(1996).
CC -1- FUNCTION: Could play a role in the determination of the epiblastic
CC cells that subsequently give rise to the mesoderm.
CC -1- TISSUE SPECIFICITY: Expressed at low level in specific organs of

```

CC the adult animal such as spleen, heart, lung and brain. During  
 CC gastrulation, expressed in the forming mesoderm. In later stages  
 CC of the developing heart, expression is restricted to the truncus  
 CC arteriosus.  
 CC -!- DEVELOPMENTAL STAGE: First expressed prior to the onset of  
 CC gastrulation (early streak stage), then continues throughout  
 CC embryonic development.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL; M87331; AAA37459.1; -; mRNA.  
 CC EMBL; X94083; CAA63827.1; -; Genomic DNA.  
 CC PIR; I49612; I49612.  
 CC HSSP; P00740; 1EDM.  
 CC Ensemble; ENSMUSG0000032494; Mus musculus.  
 CC MGI; MGI:98658; Tdglf.  
 CC GO; GO:0005615; C:extracellular space; TAS.  
 CC GO; GO:0019897; C:extrinsic to plasma membrane; IDA.  
 CC GO; GO:001763; P:branching morphogenesis; IDA.  
 CC GO; GO:0030154; P:cell differentiation; TAS.  
 CC GO; GO:0008595; P:determination of anterior/posterior axis, e. . .; IMP.  
 CC GO; GO:0007507; P:heart development; IDA.  
 CC GO; GO:0030335; P:positive regulation of cell migration; IDA.  
 CC InterPro; IPR000742; EGF\_1.  
 CC InterPro; IPR006209; EGF\_like.  
 CC Pfam; PF00008; EGF; 1.  
 CC PROSITE; PS00022; EGF\_1; 1.  
 CC PROSITE; PS50026; EGF\_3; 1.  
 CC EGF-like domain; Glycoprotein; Growth factor; Signal.  
 FT SIGNAL 1 17 Potential.  
 FT CHAIN 18 171 Teratocarcinoma-derived growth factor.  
 FT DOMAIN 62 91 EGF-like.  
 FT CARBOHYD 63 63 N-linked (GlcNAc. . .) (Potential).  
 FT DISULFID 66 73 By similarity.  
 FT DISULFID 67 79 By similarity.  
 FT DISULFID 81 90 By similarity.  
 FT DISULFID 99 117 By similarity.  
 FT DISULFID 112 133 By similarity.  
 FT DISULFID 115 124 By similarity.  
 SQ SEQUENCE 171 AA; 18754 MW; C52051AEACDB5380 CRC64;  
 Query Match 94.3%; Score 199; DB 1; Length 171;  
 Best Local Similarity 94.1%; Pred. No. 1.1e-16;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LNRTCLNGGTCMLGSCFACPPSFYGRNCEHDVR 35  
 DB 62 LNKTCLNGGTCILGSCFACPPSFYGRNCEHDVR 95  
 RESULT 6  
 QYQ06\_MOUSE  
 ID QYQ06\_MOUSE PRELIMINARY; PRT; 171 AA.  
 AC QYQ06;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Tdglf protein.  
 GN Name=Tdglf;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Blastocyst;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Alekshun S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Blastocyst;  
 RA Strausberg R.;  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC052646; AAH52646.1; -; mRNA.  
 DR HSSP; P00740; 1EDM.  
 DR MGI; MGI:98658; Tdglf.  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR GO; GO:0019897; C:extrinsic to plasma membrane; IDA.  
 DR GO; GO:0008595; P:determination of anterior/posterior axis, e. . .; IMP.  
 DR InterPro; IPR000742; EGF\_1.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR006210; IEGF.  
 DR Pfam; PF00008; EGF; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS50026; EGF\_3; 1.  
 SQ SEQUENCE 171 AA; 18654 MW; C53400EBACDB5380 CRC64;  
 Query Match 94.3%; Score 199; DB 2; Length 171;  
 Best Local Similarity 94.1%; Pred. No. 1.1e-16;  
 Matches 32; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LNRTCLNGGTCMLGSCFACPPSFYGRNCEHDVR 35  
 DB 62 LNKTCLNGGTCILGSCFACPPSFYGRNCEHDVR 95  
 RESULT 7  
 QYIAT2\_BRARE  
 ID QYIAT2\_BRARE PRELIMINARY; PRT; 66 AA.  
 AC QYIAT2;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE CripTo (Fragment).  
 GN Name=cripTo;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC Del Giacco L., Cotelli F.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF151352; AAF73188.1; -; mRNA.  
 DR ZFIN; ZDB-GENE-000710-5; cripTo.  
 FT NON\_TER 1 1  
 FT NON\_TER 66 66  
 SQ SEQUENCE 66 AA; 7275 MW; 5C6ABB16D5832871 CRC64;

```

Query Match      87.7%; Score 185; DB 2; Length 66;
Best Local Similarity 91.4%; Pred. Mismatches 3; Indels 0; Gaps 0;
Matches 32; Conservative

QY 1 ELNRTCLNGTCLMGSCFACPPSFYGRNCEHDVR 35
Db 1 ELNRTCLNGTCLMGSCFACPPSFYGRNCEHDVR 35

RESULT 8
ID O57517 BRARE PRELIMINARY; PRT; 183 AA.
AC O57517;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE One-eyed pinhead short form protein.
GN Name=oeep;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98117252; PubMed=9458048; DOI=10.1016/S0092-8674(00)80918-6;
RA Zhang J., Talbot W.S., Schier A.F.;
RT "Positional cloning identifies zebrafish one-eyed pinhead as a
RT permissive EGF-related ligand required during gastrulation.";
RL Cell 92:241-251(1998).
DR EMBL; AF041441; AAC04339.1; -; mRNA.
DR HSSP; P00750; ITPG.
DR Ensembl; ENSDARG00000035095; Danio rerio.
DR ZFIN; ZDB-GENE-990415-198; oep.
DR InterPro; IPR006209; EGF like.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
SQ SEQUENCE 183 AA; 20404 MW; 1471447894E1851B CRC64;

Query Match      69.7%; Score 147; DB 2; Length 183;
Best Local Similarity 72.7%; Pred. No. 3e-10;
Matches 24; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 NRTCCNGGTCLMGSCFACPPSFYGRNCEHDVR 35
Db 80 SRTCCNGGTCLMGSCFACPPSFYGRNCEHDVR 112

RESULT 10
OS7516 BRARE
ID O57516 BRARE PRELIMINARY; PRT; 190 AA.
AC O57516;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE One-eyed pinhead long form protein.
GN Name=oeep;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98117252; PubMed=9458048; DOI=10.1016/S0092-8674(00)80918-6;
RA Zhang J., Talbot W.S., Schier A.F.;
RT "Positional cloning identifies zebrafish one-eyed pinhead as a
RT permissive EGF-related ligand required during gastrulation.";
RL Cell 92:241-251(1998).
DR EMBL; AF041440; AAC04338.1; -; mRNA.
DR HSSP; P00750; ITPG.
DR Ensembl; ENSDARG00000035095; Danio rerio.
DR ZFIN; ZDB-GENE-990415-198; oep.
DR InterPro; IPR006209; EGF like.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
SQ SEQUENCE 190 AA; 21199 MW; 8D127CAD0625D28A CRC64;

Query Match      69.7%; Score 147; DB 2; Length 190;
Best Local Similarity 72.7%; Pred. No. 3.1e-10;
Matches 24; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 NRTCCNGGTCLMGSCFACPPSFYGRNCEHDVR 35
Db 80 SRTCCNGGTCLMGSCFACPPSFYGRNCEHDVR 112

RESULT 11
Q800J2 XENLA
Q800J2 XENLA PRELIMINARY; PRT; 190 AA.
ID Q800J2 XENLA PRELIMINARY; PRT; 190 AA.
AC Q800J2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

```

```

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE FRL-1
GN Names=FRL-1;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yabe S., Tanegashima K., Haramoto Y., Takahashi S., Fujii T.,
RA Kozuma S., Taketani Y., Asashima M.;
RT "FRL-1, a member of the EGF-CFC family, is essential for neural
RT differentiation in Xenopus early development.";
RL Development 0:0-0(2003).
DR EMBL; AB105859; BAC65863.1; -; mRNA.
DR InterPro; IPR006209; EGF like.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
SQ SEQUENCE 190 AA; 21337 MW; 6079CA6FBC1F0A83 CRC64;

Query Match 68.7%; Score 145; DB 2; Length 190;
Best Local Similarity 65.7%; Pred. No. 5.5e-10;
Matches 23; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSPFCACPPSFYGRNCEHDVR 35
:|||||:|||||:|||||:|||||:|||||:
Db 78 KLNKKCCQNGGTCFLGTFCICPKQFTGRHCHEHRR 112

#
RESULT 12
Q91649 XENLA
ID Q91649_XENLA PRELIMINARY; PRT; 190 AA.
AC Q91649;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Fibroblast growth factor receptor ligand 1.
GN Name=FRL1;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kinoshita N., Minshull J., Kirschner M.W.;
RT "The identification of two novel ligands of the FGF receptor by a
RT yeast screening method and their activity in Xenopus development.";
RL Cell 83:621-630(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96069863; PubMed=7585965; DOI=10.1016/0092-8674(95)90102-7;
RA Kinoshita N., Minshull J., Kirschner M.W.;
RT "The identification of two novel ligands of the FGF receptor by a
RT yeast screening method and their activity in Xenopus development.";
RL Cell 83:621-630(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=11024280; DOI=10.1016/S0378-1119(00)00337-1;
RA Colas J.-F., Schoenwolf G.C.;
RT "Subtractive hybridization identifies chick-cripto, a novel EGF-CFC
RT ortholog expressed during gastrulation, neurulation and early
RT cardiogenesis.";
RL Gene 255:205-217(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Colas J.-F., Schoenwolf G.C.;
RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Schlang T., Schnipkowitz I., Andree B., Ebert A., Zile M.H.,
RA Arnold H.-H., Brand T.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF228760; AAF97868.1; -; mRNA.
DR EMBL; AF228760; AAF97868.1; -; mRNA.
DR EMBL; AF228762; AAF97869.1; -; Genomic_DNA.
DR EMBL; AF228761; AAF97869.1; JOINED; Genomic_DNA.
DR HSSP; P00749; IURK.
DR Ensemble; ENSGALG0000012623; Gallus gallus.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR Pfam; PF00008; EGF; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS50026; EGF_3; 1.
KW EGF-like domain; Transmembrane.
SQ SEQUENCE 193 AA; 22332 MW; E859A98F2DE6325F CRC64;

Query Match 65.4%; Score 138; DB 2; Length 193;
Best Local Similarity 73.3%; Pred. No. 4.1e-09;
Matches 22; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 LNRTCLNGGTCMLGSPFCACPPSFYGRNCE 31
:|||||:|||||:|||||:|||||:|||||
Db 86 LNRNCCQNGGTCILGAFCAFCPCPKHFSGRHCE 115

#
RESULT 14
Q93T05 HUMAN
ID Q93T05_HUMAN PRELIMINARY; PRT; 223 AA.
AC Q93T05;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein CRYPTIC.
GN Name=CRYPTIC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Matches 23; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSPFCACPPSFYGRNCEHDVR 35
:|||||:|||||:|||||:|||||:|||||:
Db 78 KLNKKCCQNGGTCFLGTFCICPKQFTGRHCHEHRR 112

#
RESULT 13
Q91803 CHICK
ID Q91803_CHICK PRELIMINARY; PRT; 193 AA.
AC Q91803;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE CRYPTO-related factor 1 (CRYPTIC).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=11024280; DOI=10.1016/S0378-1119(00)00337-1;
RA Colas J.-F., Schoenwolf G.C.;
RT "Subtractive hybridization identifies chick-cripto, a novel EGF-CFC
RT ortholog expressed during gastrulation, neurulation and early
RT cardiogenesis.";
RL Gene 255:205-217(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Colas J.-F., Schoenwolf G.C.;
RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Schlang T., Schnipkowitz I., Andree B., Ebert A., Zile M.H.,
RA Arnold H.-H., Brand T.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF228760; AAF97868.1; -; mRNA.
DR EMBL; AF228760; AAF97868.1; -; mRNA.
DR EMBL; AF228762; AAF97869.1; -; Genomic_DNA.
DR EMBL; AF228761; AAF97869.1; JOINED; Genomic_DNA.
DR HSSP; P00749; IURK.
DR Ensemble; ENSGALG0000012623; Gallus gallus.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR Pfam; PF00008; EGF; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS50026; EGF_3; 1.
KW EGF-like domain; Transmembrane.
SQ SEQUENCE 193 AA; 22332 MW; E859A98F2DE6325F CRC64;

Query Match 65.4%; Score 138; DB 2; Length 193;
Best Local Similarity 73.3%; Pred. No. 4.1e-09;
Matches 22; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 LNRTCLNGGTCMLGSPFCACPPSFYGRNCE 31
:|||||:|||||:|||||:|||||:|||||
Db 86 LNRNCCQNGGTCILGAFCAFCPCPKHFSGRHCE 115

#
RESULT 14
Q93T05 HUMAN
ID Q93T05_HUMAN PRELIMINARY; PRT; 223 AA.
AC Q93T05;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein CRYPTIC.
GN Name=CRYPTIC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1] NUCLEOTIDE SEQUENCE.  
RA Cotton M., Du H.;  
RT "The sequence of Homo sapiens BAC clone RP11-356110.";  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Waterston R.H.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RA Waterston R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RA Wilson R.K.;  
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC013269; AY14955.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 223 AA; 24642 MW; 457EB443BC0CF1A5 CRC64;  
  
Query Match 64.5%; Score 136; DB 2; Length 223;  
Best Local Similarity 73.3%; Pred. No. 8.2e-09;  
Matches 22; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
  
QY 6 CCLNGGTCMLGSCFACPPSFYGRNCEHVDV 35  
DB 90 CCRNGGTCVLGSCFVCVPAHFTGRYCEHVDV 119  
  
RESULT 15  
ID Q9GZR3 HUMAN PRELIMINARY; PRT; 223 AA.  
AC Q9GZR3;  
DT 01-MAR-2001 (TRENBLrel. 16, Created)  
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)  
DE Cryptic.  
GN Name=CRCL;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1] NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20517351; PubMed=11062482; DOI=10.1038/91695;  
RA Bamford R.N., Roessler E., Burdine R., Saplakoglu U., Cruz J.D.,  
RA Splitt M., Towbin J., Marino B., Schier A.F., Shen M.M., Muenke M.,  
RA Casey B.;  
RT "Loss-of-function mutations in the EGF-CFC gene CRCL are associated with human left-right laterality defects.";  
RL Nat. Genet. 26:365-369(2000).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Lung; and PCR rescued clones;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshlyuki S., Carninci P., Frange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=PCR rescued clones;  
RA Director MGC Project;  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Lung;  
RA Director MGC Project;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Lung;  
RG NIH MGC Project;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF312925; AAG42475.1; -; Genomic\_DNA.  
DR EMBL; AF312769; AAG30294.1; -; mRNA.  
DR EMBL; BC059508; AAH59508.1; -; mRNA.  
DR EMBL; BC074825; AAH74825.1; -; mRNA.  
DR EMBL; BC074826; AAH74826.1; -; mRNA.  
DR HSSP; P00750; 1TPG.  
DR HGNC; HGNC:18292; CFC1.  
DR GO; GO:0007368; P:determination of left/right symmetry; NAS.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR006210; IEGF.  
DR Pfam; PF00008; EGF; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS00026; EGF\_3; 1.  
KW EGF-like domain; Transmembrane.  
SQ SEQUENCE 223 AA; 24612 MW; B52852A00ABCFLA3 CRC64;  
  
Query Match 64.5%; Score 136; DB 2; Length 223;  
Best Local Similarity 73.3%; Pred. No. 8.2e-09;  
Matches 22; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
  
QY 6 CCLNGGTCMLGSCFACPPSFYGRNCEHVDV 35  
DB 90 CCRNGGTCVLGSCFVCVPAHFTGRYCEHVDV 119

Search completed: February 4, 2006, 14:37:35  
Job time : 154.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 14:37:51 ; Search time 34.5 Seconds  
(without alignments)  
83.874 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_77\_111  
Perfect score: 211  
Sequence: 1 ELNRTCCCLNGGTCMLGSCACPPSYGRNCEHDVR 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5 COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/6 COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/H COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/PCTUS COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/RE COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211	100.0	174	6	5256643-3
2	211	100.0	188	1	US-07-749-001-3
3	211	100.0	188	1	US-08-154-198-3
4	211	100.0	188	1	US-08-463-335-3
5	211	100.0	188	1	US-08-464-023A-3
6	211	100.0	188	1	US-08-471-371-7
7	211	100.0	192	2	US-09-949-016-8959
8	211	100.0	192	2	US-09-949-016-11113
9	203	96.2	188	1	US-07-749-001-5
10	203	96.2	188	1	US-08-154-198-5
11	203	96.2	188	1	US-08-463-335-5
12	203	96.2	188	1	US-08-464-023A-5
13	203	96.2	192	2	US-09-949-016-7062
14	202	95.7	187	6	5256643-2
15	199	94.3	160	1	US-08-441-629-7
16	199	94.3	160	2	US-08-776-207-7
17	199	94.3	160	2	US-09-507-773-7
18	199	94.3	160	2	US-10-016-447-7
19	199	94.3	160	2	PCT-US95-09172-7
20	145	68.7	190	1	US-08-441-629-4
21	145	68.7	190	2	US-08-776-207-4
22	145	68.7	190	2	US-09-507-773-4
23	145	68.7	190	2	US-10-016-447-4
24	145	68.7	190	2	PCT-US95-09172-4
25	136	64.5	230	1	US-08-471-371-2
26	123	58.3	20	6	5256643-6
27	107	50.7	94	1	US-08-208-008C-14

28	105	49.8	379	2	US-09-907-794A-4	Sequence 4, Appli
29	105	49.8	379	2	US-09-905-125A-4	Sequence 4, Appli
30	105	49.8	379	2	US-09-902-775A-4	Sequence 4, Appli
31	105	49.8	379	2	US-09-906-700-4	Sequence 4, Appli
32	105	49.8	379	2	US-09-903-603A-4	Sequence 4, Appli
33	105	49.8	379	2	US-09-904-920A-4	Sequence 4, Appli
34	105	49.8	379	2	US-09-909-064-4	Sequence 4, Appli
35	105	49.8	379	2	US-09-905-381A-4	Sequence 4, Appli
36	105	49.8	379	2	US-09-906-618-4	Sequence 4, Appli
37	105	49.8	379	2	US-09-906-646-4	Sequence 4, Appli
38	105	49.8	379	2	US-09-904-462-4	Sequence 4, Appli
39	105	49.8	379	2	US-09-902-736A-4	Sequence 4, Appli
40	105	49.8	379	2	US-09-906-722A-4	Sequence 4, Appli
41	105	49.8	380	2	US-09-205-258-441	Sequence 441, App
42	105	49.8	380	2	US-10-004-860-441	Sequence 441, App
43	104	49.3	68	1	US-08-208-008C-13	Sequence 13, Appl
44	98.5	46.7	642	2	US-08-872-855-10	Sequence 10, Appl
45	96.5	45.7	2523	1	US-08-185-432-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1  
5256643-3  
; Patent No. 5256643  
; APPLICANT: Persico, Maria G.; Salomon, David S.  
; TITLE OF INVENTION: HUMAN CRYPTO PROTEIN  
; NUMBER OF SEQUENCES: 18  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/530,165  
; FILING DATE: 29-MAY-1990  
; SEQ ID NO: 3:  
; LENGTH: 174  
5256643-3

Query Match 100.0%; Score 211; DB 6; Length 174;  
Best Local Similarity 100.0%; Pred. No. 1.2e-15;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELNRTCCCLNGGTCMLGSCACPPSYGRNCEHDVR 35  
|||||  
Db 63 ELNRTCCCLNGGTCMLGSCACPPSYGRNCEHDVR 97

RESULT 2  
US-07-749-001-3  
; Sequence 3, Application US/07749001  
; Patent No. 5264557  
; GENERAL INFORMATION:  
; APPLICANT: Salomon, David S.  
; APPLICANT: Persico, Maria G.  
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/749,001  
; FILING DATE: 19910823  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Scott, Watson T.  
; REGISTRATION NUMBER: 26,581

```
; REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-749-001-3

Query Match 100.0%; Score 211; DB 1; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 35
Db 77 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 111

RESULT 3
US-08-154-198-3
; Sequence 3, Application US/08154198
; Patent No. 5620866
; GENERAL INFORMATION:
; APPLICANT: SALOMON, David S.
; APPLICANT: PERSICO, Maria G.
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/154,198
; FILING DATE: 17-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,001
; FILING DATE: 23-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-154-198-3

Query Match 100.0%; Score 211; DB 1; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 35
Db 77 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 111
```

```
RESULT 4
US-08-463-335-3
; Sequence 3, Application US/08463335
; Patent No. 5650285
; GENERAL INFORMATION:
; APPLICANT: Salomon, David S.
; APPLICANT: Persico, Maria G.
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,335
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,001
; FILING DATE: 23-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 248453 CUSH
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-335-3

Query Match 100.0%; Score 211; DB 1; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 35
Db 77 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 111

RESULT 5
US-08-464-023A-3
; Sequence 3, Application US/08464023A
; Patent No. 5854399
; GENERAL INFORMATION:
; APPLICANT: Salomon, David S.
; APPLICANT: Persico, Maria G.
; TITLE OF INVENTION: A Human Crypto-Related Gene
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



```
/
/
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/464,023A
/ FILING DATE: 05-JUN-1995
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/749,001
/ FILING DATE: 23-AUG-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/154,198
/ FILING DATE: 17-NOV-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bastian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 15280-63-1-1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 543-9600
/ TELEFAX: (415) 543-5043
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 188 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-464-023A-3

Query Match 100.0%; Score 211; DB 1; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLGGTCTMLGSCFACPPSFYGRNCEHDVR 35
Db 77 ELNRTCLGGTCTMLGSCFACPPSFYGRNCEHDVR 111

RESULT 6
US-08-471-371-7
/ Sequence 7, Application US/08471371
/ Patent No. 5981215
/ GENERAL INFORMATION:
/ APPLICANT: Weissner, Paul S.
/ APPLICANT: Coleman, Timothy A.
/ TITLE OF INVENTION: HUMAN CRIPTIN GROWTH FACTOR
/ NUMBER OF SEQUENCES: 7
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
/ STREET: 6 Becker Farm Road
/ CITY: Roseland
/ STATE: N.J.
/ COUNTRY: U.S.A
/ ZIP: 07068
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/471,371
/ FILING DATE: 06-JUN-1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ferraro, Gregory D.
/ REGISTRATION NUMBER: 36,134
/ REFERENCE/DOCKET NUMBER: 32580-455
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (201) 994-1700
/ TELEFAX: (201) 994-1744
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 188 amino acids
```

```
/
/
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-471-371-7

Query Match 100.0%; Score 211; DB 1; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLGGTCTMLGSCFACPPSFYGRNCEHDVR 35
Db 77 ELNRTCLGGTCTMLGSCFACPPSFYGRNCEHDVR 111

RESULT 7
US-09-949-016-8959
/ Sequence 8959, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 8959
/ LENGTH: 192
/ TYPE: PRT
/ ORGANISM: Human
/ US-09-949-016-8959

Query Match 100.0%; Score 211; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLGGTCTMLGSCFACPPSFYGRNCEHDVR 35
Db 81 ELNRTCLGGTCTMLGSCFACPPSFYGRNCEHDVR 115

RESULT 8
US-09-949-016-11113
/ Sequence 11113, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11113
/ LENGTH: 192
/ TYPE: PRT
/ ORGANISM: Human
/ US-09-949-016-11113
```

```
Query Match          100.0%; Score 211; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACPPSFYGRNCCHDVR 35
Db 81 ELNRTCLNGGTCMLGSCFACPPSFYGRNCCHDVR 115

RESULT 9
US-07-749-001-5
; Sequence 5, Application US/07749001
; Patent No. 5264557
; GENERAL INFORMATION:
; APPLICANT: Salomon, David S.
; APPLICANT: Persico, Maria G.
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19910823
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 248453 CUSH
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-749-001-5

Query Match          96.2%; Score 203; DB 1; Length 188;
Best Local Similarity 97.1%; Pred. No. 9.4e-15;
Matches 34; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACPPSFYGRNCCHDVR 35
Db 77 ELNRTCLNGGTCMLGSCFACPPSFYGRNCCHDVR 111

RESULT 10
US-08-154-198-5
; Sequence 5, Application US/08154198
; Patent No. 5620866
; GENERAL INFORMATION:
; APPLICANT: SALOMON, David S.
; APPLICANT: PERSICO, Maria G.
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
```

```
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/154,198
; FILING DATE: 17-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,001
; FILING DATE: 23-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 15280-63-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-154-198-5

Query Match          96.2%; Score 203; DB 1; Length 188;
Best Local Similarity 97.1%; Pred. No. 9.4e-15;
Matches 34; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACPPSFYGRNCCHDVR 35
Db 77 ELNRTCLNGGTCMLGSCFACPPSFYGRNCCHDVR 111

RESULT 11
US-08-463-335-5
; Sequence 5, Application US/08463335
; Patent No. 5650285
; GENERAL INFORMATION:
; APPLICANT: Salomon, David S.
; APPLICANT: Persico, Maria G.
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,335
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,001
; FILING DATE: 23-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
```

```
; TELEFAX: (202)822-0944
; TELEX: 248453 CUSH
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-335-5

Query Match          96.2%; Score 203; DB 1; Length 188;
Best Local Similarity 97.1%; Pred. No. 9.4e-15;
Matches 34; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVR 35
Db 77 ELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVR 111

RESULT 12
US-08-464-023A-5
; Sequence 5, Application US/08464023A
; Patent No. 5854399
; GENERAL INFORMATION:
; APPLICANT: Salomon, David S.
; APPLICANT: Persico, Maria G.
; TITLE OF INVENTION: A Human Cripto-Related Gene
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,023A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,001
; FILING DATE: 23-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/154,198
; FILING DATE: 17-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 15280-63-1-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-464-023A-5

Query Match          96.2%; Score 203; DB 1; Length 188;
Best Local Similarity 97.1%; Pred. No. 9.4e-15;
Matches 34; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVR 35
Db 77 ELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVR 111

; TELEFAX: (202)822-0944
; TELEX: 248453 CUSH
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-949-016-7062
; Sequence 7062, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7062
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7062

Query Match          96.2%; Score 203; DB 2; Length 192;
Best Local Similarity 97.1%; Pred. No. 9.6e-15;
Matches 34; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVR 35
Db 81 ELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVR 115

RESULT 14
52566643-2
; Patent No. 52566643
; APPLICANT: Persico, Maria G.; Salomon, David S.
; TITLE OF INVENTION: HUMAN CRIPTO PROTEIN
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/530,165
; FILING DATE: 29-MAY-1990
; SEQ ID NO: 2
; LENGTH: 187
52566643-2

Query Match          95.7%; Score 202; DB 6; Length 187;
Best Local Similarity 97.1%; Pred. No. 1.2e-14;
Matches 34; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVR 35
Db 77 ELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVR 111

RESULT 15
US-08-441-629-7
; Sequence 7, Application US/08441629
; Patent No. 5766923
; GENERAL INFORMATION:
; APPLICANT: Kirschner, Marc W.
; APPLICANT: Kinoshita, No. 5766923iyuki
; TITLE OF INVENTION: RECEPTOR-LIGAND ASSAY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
```

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441,629
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,217
; FILING DATE: 22-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: HU95-01A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 160 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-441-629-7

Query Match 94.3%; Score 199; DB 1; Length 160;
Best Local Similarity 94.1%; Pred. No. 2.2e-14;
Matches 32; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNRTCCLNGGTCMLGSFCACPPSFYGRNCEHDVR 35
Db ||:|||||:|||||:|||||:|||||:|||||:
62 LNKTCCLNGGTCILGSFCACPPSFYGRNCEHDVR 95

Search completed: February 4, 2006, 14:39:45
Job time : 35.5 secs
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 14:38:46 ; Search time 123.5 Seconds  
(without alignments)  
118.413 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_77\_111  
Perfect score: 211  
Sequence: 1 ELNRRCTCLNGGTCMLGSCACPPSYGRNCEHDVR 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main:  
1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep:  
2: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep:  
3: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep:  
4: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep:  
5: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep:  
6: /cgn2\_6/prodata/1/pubpaa/US11\_PUBCOMB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211	100.0	38	4	US-10-390-566-5
2	211	100.0	129	4	US-10-390-566-4
3	211	100.0	139	4	US-10-390-566-3
4	211	100.0	174	4	US-10-665-602-7
5	211	100.0	188	4	US-10-241-220-58
6	211	100.0	188	4	US-10-388-578-6
7	211	100.0	188	4	US-10-390-566-1
8	211	100.0	188	4	US-10-362-597A-3
9	211	100.0	188	4	US-10-362-597A-6
10	211	100.0	188	4	US-10-362-597A-96
11	211	100.0	188	4	US-10-407-481-3
12	211	100.0	188	4	US-10-407-481-6
13	211	100.0	188	4	US-10-407-481-96
14	211	100.0	188	4	US-10-712-124-58
15	211	100.0	188	4	US-10-693-538-1
16	211	100.0	188	4	US-10-389-431-6
17	211	100.0	188	4	US-10-816-476-3
18	211	100.0	188	4	US-10-816-476-6
19	211	100.0	188	4	US-10-816-476-96
20	211	100.0	188	5	US-10-872-972-58
21	211	100.0	188	5	US-10-872-991-58
22	211	100.0	188	5	US-10-491-997-28
23	211	100.0	188	5	US-10-940-431-2
24	211	100.0	188	5	US-10-983-340-13
25	211	100.0	266	4	US-10-390-566-8
26	211	100.0	360	4	US-10-390-566-7
27	211	100.0	367	4	US-10-390-566-6

```

28 206 97.6 38 4 US-10-390-566-12 Sequence 12, Appl
29 206 97.6 129 4 US-10-390-566-11 Sequence 11, Appl
30 206 97.6 139 4 US-10-390-566-10 Sequence 10, Appl
31 206 97.6 188 4 US-10-390-566-9 Sequence 9, Appl
32 206 97.6 266 4 US-10-390-566-15 Sequence 15, Appl
33 206 97.6 360 4 US-10-390-566-14 Sequence 14, Appl
34 206 97.6 367 4 US-10-390-566-13 Sequence 13, Appl
35 204 96.7 37 4 US-10-470-013-2 Sequence 2, Appl
36 203 96.2 38 4 US-10-390-566-18 Sequence 18, Appl
37 203 96.2 129 4 US-10-390-566-17 Sequence 17, Appl
38 203 96.2 139 4 US-10-390-566-16 Sequence 16, Appl
39 203 96.2 139 4 US-10-390-566-23 Sequence 23, Appl
40 203 96.2 188 4 US-10-390-566-2 Sequence 2, Appl
41 203 96.2 188 4 US-10-362-597A-4 Sequence 4, Appl
42 203 96.2 188 4 US-10-407-481-4 Sequence 4, Appl
43 203 96.2 188 4 US-10-693-538-2 Sequence 2, Appl
44 203 96.2 188 4 US-10-816-476-4 Sequence 4, Appl
45 203 96.2 188 5 US-10-491-997-72 Sequence 72, Appl

```

## ALIGNMENTS

```

RESULT 1
US-10-390-566-5
; Sequence 5, Application US/10390566
; Publication NO. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-5

Query Match 100.0%; Score 211; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 3.5e-16;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRRCTCLNGGTCMLGSCACPPSYGRNCEHDVR 35
Db 3 ELNRRCTCLNGGTCMLGSCACPPSYGRNCEHDVR 37

RESULT 2
US-10-390-566-4
; Sequence 4, Application US/10390566
; Publication NO. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148

```

;  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 129  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-4

Query Match 100.0%; Score 211; DB 4; Length 129;  
Best Local Similarity 100.0%; Pred. No. 9.6e-16;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSFCACPPSFYGRNCEHDVR 35  
|||  
DB 37 ELNRTCLNGGTCMLGSFCACPPSFYGRNCEHDVR 71

RESULT 3  
US-10-390-566-3  
; Sequence 3, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-3

Query Match 100.0%; Score 211; DB 4; Length 139;  
Best Local Similarity 100.0%; Pred. No. 1e-15;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSFCACPPSFYGRNCEHDVR 35  
|||  
DB 47 ELNRTCLNGGTCMLGSFCACPPSFYGRNCEHDVR 81

RESULT 4  
US-10-665-602-7  
; Sequence 7, Application US/10665602  
; Publication No. US20040086967A1  
; GENERAL INFORMATION:  
; APPLICANT: Meissner, Paul S.  
; TITLE OF INVENTION: Human Cripitin Growth Factor  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: MD  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/665,602  
FILING DATE: 22-Sep-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/393,023A  
FILING DATE: 09-SEP-1999  
APPLICATION NUMBER: US 08/471,371  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Marks, Michelle S.  
REGISTRATION NUMBER: 41,971  
REFERENCE/DOCKET NUMBER: PF2000D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-665-602-7  
Query Match 100.0%; Score 211; DB 4; Length 174;  
Best Local Similarity 100.0%; Pred. No. 1.2e-15;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSFCACPPSFYGRNCEHDVR 35  
|||  
DB 63 ELNRTCLNGGTCMLGSFCACPPSFYGRNCEHDVR 97

RESULT 5  
US-10-241-220-58  
; Sequence 58, Application US/10241220  
; Publication No. US20030148408A1  
; GENERAL INFORMATION:  
; APPLICANT: Frantcz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P5010R1-US  
; CURRENT APPLICATION NUMBER: US/10/241,220  
; CURRENT FILING DATE: 2002-12-13  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 58  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-241-220-58

Query Match 100.0%; Score 211; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.3e-15;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSFCACPPSFYGRNCEHDVR 35  
|||  
DB 77 ELNRTCLNGGTCMLGSFCACPPSFYGRNCEHDVR 111

RESULT 6  
US-10-241-220-58  
; Sequence 58, Application US/10241220  
; Publication No. US20030148408A1  
; GENERAL INFORMATION:  
; APPLICANT: Frantcz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P5010R1-US  
; CURRENT APPLICATION NUMBER: US/10/241,220  
; CURRENT FILING DATE: 2002-12-13  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 58  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-241-220-58

Query Match 100.0%; Score 211; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.3e-15;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSFCACPPSFYGRNCEHDVR 35  
|||  
DB 77 ELNRTCLNGGTCMLGSFCACPPSFYGRNCEHDVR 111

RESULT 6

```

US-10-388-578-6
; Sequence 6, Application US/10388578
; Publication No. US20030224411A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence
; APPLICANT: Geron Corporation
; APPLICANT: Ralph, Brandenberger
; APPLICANT: Joseph, Gold D.
; APPLICANT: John, Irving
; APPLICANT: Mandalam, Ramkumar
; APPLICANT: Mok, Michael
; APPLICANT: Shelton, Dawne
; TITLE OF INVENTION: Genes that are Up- or Down-Regulated During Differentiation of Hu
; FILE OF INVENTION: Embryonic Stem Cells
; FILE REFERENCE: 135/001
; CURRENT APPLICATION NUMBER: US/10/388,578
; CURRENT FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: Custom
; SEQ ID NO 6
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-388-578-6

Query Match      100.0%; Score 211; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ELNRTCLNGGTCTMLGSCFACPPSFYGRNCEHDVR 35
DB      77  ELNRTCLNGGTCTMLGSCFACPPSFYGRNCEHDVR 111

RESULT 7
US-10-390-566-1
; Sequence 1, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Doman, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-1

Query Match      100.0%; Score 211; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ELNRTCLNGGTCTMLGSCFACPPSFYGRNCEHDVR 35
DB      77  ELNRTCLNGGTCTMLGSCFACPPSFYGRNCEHDVR 111

RESULT 8
US-10-362-597A-3
; Sequence 3, Application US/10362597A
; Publication No. US20040054142A1

```

```
; PRIOR APPLICATION NUMBER: PCT/EP01/0964
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 188
; TYPE: PRT
; ORGANISM: human
US-10-362-597A-96

Query Match
Best Local Similarity 100.0%; Score 211; DB 4; Length 188;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 35
Db 77 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 111

RESULT 11
US-10-407-481-3
; Sequence 3, Application US/10407481
; Publication No. US20040138112A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237-1
; CURRENT APPLICATION NUMBER: US/10/407,481
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: PCT/EP01/09646
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Human
US-10-407-481-3

Query Match
Best Local Similarity 100.0%; Score 211; DB 4; Length 188;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 35
Db 77 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 111

RESULT 12
US-10-407-481-6
; Sequence 6, Application US/10407481
; Publication No. US20040138112A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237-1
; CURRENT APPLICATION NUMBER: US/10/407,481
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: PCT/EP01/09646
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 188
; TYPE: PRT
; ORGANISM: human
US-10-407-481-6

Query Match
Best Local Similarity 100.0%; Score 211; DB 4; Length 188;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 35
Db 77 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 111

RESULT 13
US-10-407-481-96
; Sequence 96, Application US/10407481
; Publication No. US20040138112A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237-1
; CURRENT APPLICATION NUMBER: US/10/407,481
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: PCT/EP01/09646
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 188
; TYPE: PRT
; ORGANISM: human
US-10-407-481-96

Query Match
Best Local Similarity 100.0%; Score 211; DB 4; Length 188;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 35
Db 77 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 111

RESULT 14
US-10-712-124-58
; Sequence 58, Application US/10712124
; Publication No. US20040146907A1
; GENERAL INFORMATION:
; APPLICANT: SMITH, VICTORIA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING DYSPLASIA
; FILE REFERENCE: P2000R1
; CURRENT APPLICATION NUMBER: US/10/712,124
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 60/425,813
; PRIOR FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 123
; SEQ ID NO 58
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-712-124-58

Query Match
Best Local Similarity 100.0%; Score 211; DB 4; Length 188;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 35
Db 77 ELNRTCLNGGTCMLGSCFACACPPSFYGRNCEHDVR 111

RESULT 15
US-10-693-538-1
; Sequence 1, Application US/10693538
```



```
; Publication No. US20040146940A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Sanicola-Nadel, Michele
; APPLICANT: Williams, Kevin
; APPLICANT: Schiffer, Susan
; APPLICANT: Rayhorn, Paul
; TITLE OF INVENTION: Crypto Blocking Antibodies and Uses
; FILE REFERENCE: Thereof
; FILE REFERENCE: BIN-A117CN
; CURRENT APPLICATION NUMBER: US/10/693,538
; CURRENT FILING DATE: 2003-10-23
; PRIOR APPLICATION NUMBER: PCT/US02/11950
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: 60/286,782
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/293,020
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/301,091
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/367,002
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-693-538-1
```

```
Query Match      100.0%; Score 211; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ELNRTCLNGGTCMLGSCACPPSFYGRNCEHDVR 35
Db      77 ELNRTCLNGGTCMLGSCACPPSFYGRNCEHDVR 111
```

Search completed: February 4, 2006, 14:43:57  
Job time : 124.5 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 14:40:02 ; Search time 9.5 Seconds  
(without alignments)  
43.172 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_77\_111  
Perfect score: 211  
Sequence: 1 ELNRTCCLLNGGTCMLGSCFACPPSPFYGRNCHDVR 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA New:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB\_PEP.\*  
2: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB\_PEP.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB\_PEP.\*  
4: /cgn2\_6/ptodata/2/pubpaa/PT\_NEW\_PUB\_PEP.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB\_PEP.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB\_PEP.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB\_PEP.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB\_PEP.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	211	100.0	188	US-10-945-853-1	Sequence 1, Appli
2	203	96.2	188	US-10-945-853-2	Sequence 2, Appli
3	160	75.8	29	US-10-945-853-8	Sequence 8, Appli
4	105	49.8	335	US-11-136-619-30	Sequence 30, Appli
5	105	49.8	351	US-11-136-619-14	Sequence 14, Appli
6	105	49.8	379	US-10-131-826A-308	Sequence 308, Appli
7	105	49.8	379	US-11-136-619-2	Sequence 2, Appli
8	105	49.8	379	US-11-136-619-22	Sequence 22, Appli
9	105	49.8	379	US-11-136-619-23	Sequence 23, Appli
10	96	45.5	374	US-11-136-619-24	Sequence 24, Appli
11	96	45.5	378	US-11-136-619-25	Sequence 25, Appli
12	94	44.5	17	US-10-945-853-4	Sequence 4, Appli
13	92.5	43.8	816	US-11-090-439-48	Sequence 48, Appli
14	91.5	43.4	2556	US-11-050-346-67	Sequence 67, Appli
15	91	43.1	1198	US-10-453-372-880	Sequence 880, Appli
16	91	43.1	1398	US-10-055-877-46	Sequence 46, Appli
17	91	43.1	1398	US-10-453-372-872	Sequence 872, Appli
18	91	43.1	1450	US-10-055-877-48	Sequence 48, Appli
19	91	43.1	1450	US-10-453-372-874	Sequence 874, Appli
20	91	43.1	1594	US-10-453-372-860	Sequence 860, Appli
21	90.5	42.9	685	US-10-131-826A-88	Sequence 88, Appli
22	90.5	42.9	685	US-11-078-735-19	Sequence 19, Appli
23	90.5	42.9	685	US-11-050-346-64	Sequence 64, Appli
24	90.5	42.9	685	US-11-103-077-19	Sequence 19, Appli
25	89	42.2	618	US-11-078-735-18	Sequence 18, Appli

26	89	42.2	618	7	US-11-050-346-63	Sequence 63, Appli
27	89	42.2	618	7	US-11-103-077-18	Sequence 18, Appli
28	88	41.7	589	7	US-11-183-136-42	Sequence 42, Appli
29	88	41.7	1418	6	US-10-453-372-864	Sequence 864, Appli
30	87.5	41.5	954	6	US-10-453-372-1150	Sequence 1150, Appli
31	87.5	41.5	1268	6	US-10-453-372-1144	Sequence 1144, Appli
32	87.5	41.5	1268	6	US-10-453-372-1154	Sequence 1154, Appli
33	87.5	41.5	1288	6	US-10-453-372-1146	Sequence 1146, Appli
34	87.5	41.5	1288	6	US-10-453-372-1152	Sequence 1152, Appli
35	87.5	41.5	1408	6	US-10-453-372-1148	Sequence 1148, Appli
36	87.5	41.5	4913	6	US-10-453-372-1142	Sequence 1142, Appli
37	87.5	41.5	4961	6	US-10-453-372-1132	Sequence 1132, Appli
38	87	41.2	2471	7	US-11-050-346-68	Sequence 68, Appli
39	85.5	40.5	484	7	US-11-078-735-43	Sequence 43, Appli
40	85.5	40.5	484	7	US-11-050-346-37	Sequence 37, Appli
41	85.5	40.5	723	6	US-10-131-826A-346	Sequence 346, Appli
42	85.5	40.5	723	6	US-11-078-735-17	Sequence 17, Appli
43	85.5	40.5	723	7	US-11-050-346-62	Sequence 62, Appli
44	85.5	40.5	723	7	US-11-103-077-17	Sequence 17, Appli
45	85.5	40.5	864	7	US-11-050-346-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1  
US-10-945-853-1  
; Sequence 1, Application US/10945853  
; Publication No. US20050255117A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc.  
; APPLICANT: Sanicola-Nadel, Michele  
; APPLICANT: Adkins, Heather  
; APPLICANT: Miklasz, Steven Donald  
; APPLICANT: Rayhorn, Paul  
; APPLICANT: Schiffer, Susan Gail  
; APPLICANT: Williams, Kevin  
; TITLE OF INVENTION: Crypto-Specific Antibodies  
; FILE REFERENCE: BGNAL17CPCCN  
; CURRENT APPLICATION NUMBER: US/10/945,853  
; CURRENT FILING DATE: 2004-09-20  
; PRIOR APPLICATION NUMBER: PCT/US02/31462  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: PCT/US02/11950  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/286,782  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: 60/293,020  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/301,091  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/367,002  
; PRIOR FILING DATE: 2002-03-22  
; NUMBER OF SEQ ID NOS: 9  
; SEQ ID NO 1  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-945-853-1

Query Match 100.0%; Score 211; DB 6; Length 188;  
Best Local Similarity 100.0%; Pred. No. 3.6e-19;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELNRTCCLLNGGTCMLGSCFACPPSPFYGRNCHDVR 35  
|||||  
Db 77 ELNRTCCLLNGGTCMLGSCFACPPSPFYGRNCHDVR 111  
|||||

RESULT 2  
US-10-945-853-2  
; Sequence 2, Application US/10945853  
; Publication No. US20050255117A1

```
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Sanicola-Nadel, Michele
; APPLICANT: Adkins, Heather
; APPLICANT: Miklasz, Steven Donald
; APPLICANT: Rayhorn, Paul
; APPLICANT: Schiffer, Susan Gail
; APPLICANT: Williams, Kevin
; TITLE OF INVENTION: Crypto-Specific Antibodies
; FILE REFERENCE: BGN117CPCCN
; CURRENT APPLICATION NUMBER: US/10/945,853
; CURRENT FILING DATE: 2004-09-20
; PRIOR APPLICATION NUMBER: PCT/US02/31462
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: PCT/US02/11950
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/286,782
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/293,020
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/301,091
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/367,002
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 2
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-945-853-2
```

```
Query Match 96.2%; Score 203; DB 6; Length 188;
Best Local Similarity 97.1%; Pred. No. 3.3e-18;
Matches 34; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELNRTCCLLGGTCMLGSFCACPPSPFYGRNCEHDVR 35
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77 ELNRTCCLLGGTCMLGSFCACPPSPFYGRNCEHDVR 111
```

```
RESULT 3
US-10-945-853-8
; Sequence 8, Application US/10945853
; Publication NO. US20050255117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Sanicola-Nadel, Michele
; APPLICANT: Adkins, Heather
; APPLICANT: Miklasz, Steven Donald
; APPLICANT: Rayhorn, Paul
; APPLICANT: Schiffer, Susan Gail
; APPLICANT: Williams, Kevin
; TITLE OF INVENTION: Crypto-Specific Antibodies
; FILE REFERENCE: BGN117CPCCN
; CURRENT APPLICATION NUMBER: US/10/945,853
; CURRENT FILING DATE: 2004-09-20
; PRIOR APPLICATION NUMBER: PCT/US02/31462
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: PCT/US02/11950
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/286,782
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/293,020
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/301,091
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/367,002
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 8
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: Mod_res
; LOCATION: 1
; OTHER INFORMATION: N-terminal acetylation
US-10-945-853-8

Query Match 75.8%; Score 160; DB 6; Length 29;
Best Local Similarity 96.4%; Pred. No. 1.2e-13;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 LNEGTCMLGSFCACPPSPFYGRNCEHDVR 35
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 LNEGTCMLGSFCACPPSPFYGRNCEHDVR 28

RESULT 4
US-11-136-619-30
; Sequence 30, Application US/11136619
; Publication NO. US20050288226A1
; GENERAL INFORMATION:
; APPLICANT: You, Liang
; APPLICANT: He, Biao
; APPLICANT: Xu, Zhidong
; APPLICANT: Jablons, David M.
; TITLE OF INVENTION: METHODS FOR TREATING AND DIAGNOSING CANCER WITH WNT INHIBITORY
; TITLE OF INVENTION: FACTOR-1 (WIF-1)
; FILE REFERENCE: 023070-148711US
; CURRENT APPLICATION NUMBER: US/11/136,619
; CURRENT FILING DATE: 2005-05-23
; PRIOR APPLICATION NUMBER: US 60/664,241
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/573,197
; PRIOR FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 30
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-136-619-30
```

```
Query Match 49.8%; Score 105; DB 7; Length 335;
Best Local Similarity 60.0%; Pred. No. 4e-06;
Matches 15; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 7 CLNGTCMLGSFCACPPSPFYGRNCE 31
      ||||| : ||||| ||||| :
Db 190 CMNGGLCVTPGFCICPPGFYGVNCD 214

RESULT 5
US-11-136-619-14
; Sequence 14, Application US/11136619
; Publication NO. US20050288226A1
; GENERAL INFORMATION:
; APPLICANT: You, Liang
; APPLICANT: He, Biao
; APPLICANT: Xu, Zhidong
; APPLICANT: Jablons, David M.
; TITLE OF INVENTION: METHODS FOR TREATING AND DIAGNOSING CANCER WITH WNT INHIBITORY
; TITLE OF INVENTION: FACTOR-1 (WIF-1)
; FILE REFERENCE: 023070-148711US
; CURRENT APPLICATION NUMBER: US/11/136,619
; CURRENT FILING DATE: 2005-05-23
; PRIOR APPLICATION NUMBER: US 60/664,241
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/573,197
; PRIOR FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 14
; LENGTH: 351
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-11-136-619-14

Query Match          49.8%; Score 105; DB 7; Length 351;
Best Local Similarity 60.0%; Pred. No. 4.4e-06;
Matches 15; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 7 CLNGGTCMLGSCACPPSPFYGRNCE 31
Db 190 CMNGGLCVTPGFCICPPGFGVNCNCD 214

RESULT 6
US-10-131-826A-308
; Sequence 308, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 308
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-308

Query Match          49.8%; Score 105; DB 6; Length 379;
Best Local Similarity 60.0%; Pred. No. 4.4e-06;
Matches 15; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 7 CLNGGTCMLGSCACPPSPFYGRNCE 31
Db 190 CMNGGLCVTPGFCICPPGFGVNCNCD 214

; ORGANISM: Homo sapiens
US-11-136-619-22

Query Match          49.8%; Score 105; DB 7; Length 379;
Best Local Similarity 60.0%; Pred. No. 4.4e-06;
Matches 15; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 7 CLNGGTCMLGSCACPPSPFYGRNCE 31
Db 218 CMNGGLCVTPGFCICPPGFGVNCNCD 242

RESULT 7
US-11-136-619-22
; Sequence 2, Application US/11136619
; Publication No. US20050288226A1
; GENERAL INFORMATION:
; APPLICANT: You, Liang
; APPLICANT: He, Biao
; APPLICANT: Xu, Zhidong
; APPLICANT: Jablons, David M.
; TITLE OF INVENTION: METHODS FOR TREATING AND DIAGNOSING CANCER WITH WNT INHIBITORY
; FILE REFERENCE: 023070-148711US
; CURRENT APPLICATION NUMBER: US/11/136,619
; CURRENT FILING DATE: 2005-05-23
; PRIOR APPLICATION NUMBER: US 60/664,241
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/573,197
; PRIOR FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-136-619-22

Query Match          49.8%; Score 105; DB 7; Length 379;
Best Local Similarity 60.0%; Pred. No. 4.4e-06;
Matches 15; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 7 CLNGGTCMLGSCACPPSPFYGRNCE 31
Db 218 CMNGGLCVTPGFCICPPGFGVNCNCD 242

RESULT 8
US-11-136-619-22
; Sequence 22, Application US/11136619
; Publication No. US20050288226A1
; GENERAL INFORMATION:
; APPLICANT: You, Liang
; APPLICANT: He, Biao
; APPLICANT: Xu, Zhidong
; APPLICANT: Jablons, David M.
; TITLE OF INVENTION: METHODS FOR TREATING AND DIAGNOSING CANCER WITH WNT INHIBITORY
; FILE REFERENCE: 023070-148711US
; CURRENT APPLICATION NUMBER: US/11/136,619
; CURRENT FILING DATE: 2005-05-23
; PRIOR APPLICATION NUMBER: US 60/664,241
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/573,197
; PRIOR FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-136-619-22

Query Match          49.8%; Score 105; DB 7; Length 379;
Best Local Similarity 60.0%; Pred. No. 4.4e-06;
Matches 15; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 7 CLNGGTCMLGSCACPPSPFYGRNCE 31
Db 218 CMNGGLCVTPGFCICPPGFGVNCNCD 242
```

```
RESULT 9
US-11-136-619-23
; Sequence 23, Application US/11136619
; Publication No. US20050288226A1
; GENERAL INFORMATION:
; APPLICANT: You, Liang
; APPLICANT: He, Biao
; APPLICANT: Xu, Zhidong
; APPLICANT: Jablons, David M.
; TITLE OF INVENTION: METHODS FOR TREATING AND DIAGNOSING CANCER WITH WNT INHIBITORY
; FILE REFERENCE: 023070-148711US
; CURRENT APPLICATION NUMBER: US/11/136,619
; CURRENT FILING DATE: 2005-05-23
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/664,241
; PRIOR FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rattus rattus
US-11-136-619-23

Query Match      49.8%; Score 105; DB 7; Length 379;
Best Local Similarity 60.0%; Pred. No. 4.4e-06;
Matches 15; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 7 CLNGGTCMLGSCFACPPSPYGRNCE 31
Db 218 CMNGGLCVTPGFCICPPGFYGVNCD 242

RESULT 10
US-11-136-619-24
; Sequence 24, Application US/11136619
; Publication No. US20050288226A1
; GENERAL INFORMATION:
; APPLICANT: You, Liang
; APPLICANT: He, Biao
; APPLICANT: Xu, Zhidong
; APPLICANT: Jablons, David M.
; TITLE OF INVENTION: METHODS FOR TREATING AND DIAGNOSING CANCER WITH WNT INHIBITORY
; FILE REFERENCE: 023070-148711US
; CURRENT APPLICATION NUMBER: US/11/136,619
; CURRENT FILING DATE: 2005-05-23
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/664,241
; PRIOR FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Xenopus sp.
US-11-136-619-24

Query Match      45.5%; Score 96; DB 7; Length 374;
Best Local Similarity 52.0%; Pred. No. 5.4e-05;
Matches 13; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 7 CLNGGTCMLGSCFACPPSPYGRNCE 31
Db 213 CMNGGLCVTPGLICPPGFYGINCD 237

RESULT 11
US-11-136-619-25
; Sequence 25, Application US/11136619
```

```
; Publication No. US20050288226A1
; GENERAL INFORMATION:
; APPLICANT: You, Liang
; APPLICANT: He, Biao
; APPLICANT: Xu, Zhidong
; APPLICANT: Jablons, David M.
; TITLE OF INVENTION: METHODS FOR TREATING AND DIAGNOSING CANCER WITH WNT INHIBITORY
; FILE REFERENCE: 023070-148711US
; CURRENT APPLICATION NUMBER: US/11/136,619
; CURRENT FILING DATE: 2005-05-23
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/664,241
; PRIOR FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Danio rerio
US-11-136-619-25

Query Match      45.5%; Score 96; DB 7; Length 378;
Best Local Similarity 56.0%; Pred. No. 5.5e-05;
Matches 14; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 7 CLNGGTCMLGSCFACPPSPYGRNCE 31
Db 216 CLNGGLCMSPGVICPPGFYGSCE 240

RESULT 12
US-10-945-853-4
; Sequence 4, Application US/10945853
; Publication No. US20050255117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Sanicola-Nadel, Michele
; APPLICANT: Adkins, Heather
; APPLICANT: Miklasz, Steven Donald
; APPLICANT: Rayhorn, Paul
; APPLICANT: Schiffer, Susan Gail
; APPLICANT: Williams, Kevin
; TITLE OF INVENTION: Crypto-Specific Antibodies
; FILE REFERENCE: BGNA117CPCCN
; CURRENT APPLICATION NUMBER: US/10/945,853
; CURRENT FILING DATE: 2004-09-20
; PRIOR FILING DATE: PCT/US02/31462
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: PCT/US02/11950
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/286,782
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/293,020
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/301,091
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/367,002
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 4
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-945-853-4

Query Match      44.5%; Score 94; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 CPPSFYGRNCEHDVR 35
|||||
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (891)
; OTHER INFORMATION: Variable amino acid
US-11-050-346-67

Query Match      43.4%; Score 91.5; DB 7; Length 2556;
Best Local Similarity 45.9%; Pred. No. 0.00091;
Matches 17; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY      1 ELNRTCTCLNGGTG--LGSF-CACPPSFYGRNCEHDV 34
          : : : : : : : : : : : : : : : : : :
DB      987 DCTESSCFNGTCTVDGINSFTCLCPGFTGSCQHV 1023
          : : : : : : : : : : : : : : : : : :

RESULT 15
US-10-453-372-880
; Sequence 880, Application US/10453372
; Publication No. US200600323A1
; GENERAL INFORMATION:
; APPLICANT: Alcobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, A
; FILE REFERENCES: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1809
; SOFTWARE: Curasequelist version 0.1
; SEQ ID NO 880
; LENGTH: 1198
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-880

Query Match      43.1%; Score 91; DB 6; Length 1198;
Best Local Similarity 44.4%; Pred. No. 0.00056;
Matches 16; Conservative 6; Mismatches 12; Indels 2; Gaps 1;

QY      1 ELNRTCTCLNGGTGMLGSF--CACPPSFYGRNCEHDV 34
          | : : : : : : : : : : : : : : : :
DB      123 ECSASLCFHGRCVPGSAQPCPCFPFGQPRCQDV 158
          | : : : : : : : : : : : : : : : :

Search completed: February 4, 2006, 14:44:22
Job time : 9.5 secs

```

**THIS PAGE BLANK (USPTO)**



Result No.	Score	Query Match	Length	DB	ID	Description	
1	235	100.0	94	1	US-08-208-008C-14	Sequence 14, Appl	
2	235	100.0	174	6	5256643-2	Patent No. 5256643	
3	235	100.0	187	6	5256643-3	Patent No. 5256643	
4	235	100.0	188	1	US-07-749-001-3	Sequence 3, Appl	
5	235	100.0	188	1	US-07-749-001-5	Sequence 5, Appl	
6	235	100.0	188	1	US-08-154-198-3	Sequence 3, Appl	
7	235	100.0	188	1	US-08-154-198-5	Sequence 5, Appl	
8	235	100.0	188	1	US-08-463-335-3	Sequence 3, Appl	
9	235	100.0	188	1	US-08-463-335-5	Sequence 5, Appl	
10	235	100.0	188	1	US-08-464-023A-3	Sequence 3, Appl	
11	235	100.0	188	1	US-08-464-023A-5	Sequence 5, Appl	
12	235	100.0	188	1	US-08-471-371-7	Sequence 7, Appl	
13	235	100.0	192	2	US-09-949-016-7062	Sequence 7062, Ap	
14	235	100.0	192	2	US-09-949-016-8959	Sequence 8959, Ap	
15	235	100.0	192	2	US-09-949-016-1113	Sequence 1113, A	
16	194	82.6	160	1	US-08-441-629-7	Sequence 7, Appl	
17	194	82.6	160	2	US-08-776-207-7	Sequence 7, Appl	
18	194	82.6	160	2	US-09-507-773-7	Sequence 7, Appl	
19	194	82.6	160	2	US-10-016-447-7	Sequence 7, Appl	
20	194	82.6	160	4	PCR-US95-09172-7	Sequence 7, Appl	
21	88.5	37.7	190	1	US-08-441-629-4	Sequence 4, Appl	
22	88.5	37.7	190	2	US-08-776-207-4	Sequence 4, Appl	
23	88.5	37.7	190	2	US-09-507-773-4	Sequence 4, Appl	
24	88.5	37.7	190	2	US-10-016-447-4	Sequence 4, Appl	
25	88.5	37.7	190	4	PCR-US95-09172-4	Sequence 4, Appl	
26	88	37.4	230	1	US-08-471-371-2	Sequence 2, Appl	
27	65	27.7	176	1	US-08-726-525-2	Sequence 2, Appl	

Db 20 NCGSVPHDTWLPKKSLCKCKWHGQLRCFPQAFPLPGCD 56  
|||||

## RESULT 2

5256643-3  
;PATENT No. 5256643  
; APPLICANT: Persico, Maria G.; Salomon, David S.  
; TITLE OF INVENTION: HUMAN CRYPTO PROTEIN  
; NUMBER OF SEQUENCES: 18  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/530,165  
; FILING DATE: 29-MAY-1990  
; SEQ ID NO:3:  
; LENGTH: 174  
5256643-3

Query Match 100.0%; Score 235; DB 6; Length 174;  
Best Local Similarity 100.0%; Pred. No. 4.9e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKKSLCKCKWHGQLRCFPQAFPLPGCD 37  
|||||

Db 100 NCGSVPHDTWLPKKSLCKCKWHGQLRCFPQAFPLPGCD 136  
|||||

## RESULT 3

5256643-2  
;PATENT No. 5256643  
; APPLICANT: Persico, Maria G.; Salomon, David S.  
; TITLE OF INVENTION: HUMAN CRYPTO PROTEIN  
; NUMBER OF SEQUENCES: 18  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/530,165  
; FILING DATE: 29-MAY-1990  
; SEQ ID NO:2:  
; LENGTH: 187  
5256643-2

Query Match 100.0%; Score 235; DB 6; Length 187;  
Best Local Similarity 100.0%; Pred. No. 5.3e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKKSLCKCKWHGQLRCFPQAFPLPGCD 37  
|||||

Db 114 NCGSVPHDTWLPKKSLCKCKWHGQLRCFPQAFPLPGCD 150  
|||||

## RESULT 4

US-07-749-001-3  
; Sequence 3, Application US/07749001  
; Patent No. 5264557  
; GENERAL INFORMATION:  
; APPLICANT: Salomon, David S.  
; APPLICANT: Persico, Maria G.  
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/749,001  
; FILING DATE: 19910823  
; CLASSIFICATION: 435

Query Match 100.0%; Score 235; DB 6; Length 187;  
Best Local Similarity 100.0%; Pred. No. 5.3e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; ATTORNEY/AGENT INFORMATION:  
; NAME: Scott, Watson T.  
; REGISTRATION NUMBER: 26,581  
; REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 248453 CUSH  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 188 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-749-001-3

Query Match 100.0%; Score 235; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 5.3e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKKSLCKCKWHGQLRCFPQAFPLPGCD 37  
|||||

Db 114 NCGSVPHDTWLPKKSLCKCKWHGQLRCFPQAFPLPGCD 150  
|||||

## RESULT 5

US-07-749-001-5  
; Sequence 5, Application US/07749001  
; Patent No. 5264557  
; GENERAL INFORMATION:  
; APPLICANT: Salomon, David S.  
; APPLICANT: Persico, Maria G.  
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/749,001  
; FILING DATE: 19910823  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Scott, Watson T.  
; REGISTRATION NUMBER: 26,581  
; REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 248453 CUSH  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 188 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-749-001-5

Query Match 100.0%; Score 235; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 5.3e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKKSLCKCKWHGQLRCFPQAFPLPGCD 37  
|||||

Db 114 NCGSVPHDTWLPKKSLCKCKWHGQLRCFPQAFPLPGCD 150  
|||||

RESULT 6  
US-08-154-198-3  
; Sequence 3, Application US/08154198  
; Patent No. 5620866  
; GENERAL INFORMATION:  
; APPLICANT: SALOMON, David S.  
; APPLICANT: PERSICO, Maria G.  
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: Stewart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,001  
; FILING DATE: 23-AUG-1991  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 15280-63-1  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 188 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-154-198-3  
Query Match 100.0%; Score 235; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 5.3e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFLEPGCD 37  
Db 114 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFLEPGCD 150  
  
RESULT 7  
US-08-154-198-5  
; Sequence 5, Application US/08154198  
; Patent No. 5620866  
; GENERAL INFORMATION:  
; APPLICANT: SALOMON, David S.  
; APPLICANT: PERSICO, Maria G.  
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: Stewart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,001  
; FILING DATE: 23-AUG-1991  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 15280-63-1  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 188 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-154-198-5  
Query Match 100.0%; Score 235; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 5.3e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFLEPGCD 37  
Db 114 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFLEPGCD 150

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/154,198  
FILING DATE: 17-NOV-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,001  
FILING DATE: 23-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 15280-63-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-154-198-5  
Query Match 100.0%; Score 235; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 5.3e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFLEPGCD 37  
Db 114 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFLEPGCD 150  
  
RESULT 8  
US-08-463-335-3  
; Sequence 3, Application US/08463335  
; Patent No. 5650285  
; GENERAL INFORMATION:  
; APPLICANT: Salomon, David S.  
; APPLICANT: Persico, Maria G.  
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,335  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,001  
; FILING DATE: 23-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Scott, Watson T.  
; REGISTRATION NUMBER: 26,581  
; REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 248453 CUSH  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 188 amino acids

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-335-3

Query Match 100.0%; Score 235; DB 1; Length 188;
Best Local Similarity 100.0%; Pred. No. 5.3e-23;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37
Db 114 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 150

RESULT 9
US-08-463-335-5
; Sequence 5, Application US/08463335
; Patent No. 5650285
; GENERAL INFORMATION:
; APPLICANT: Salomon, David S.
; APPLICANT: Persico, Maria G.
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARY & CUSHMAN
; STREET: 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463.335
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,001
; FILING DATE: 23-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 248453 CUSH
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-335-5

Query Match 100.0%; Score 235; DB 1; Length 188;
Best Local Similarity 100.0%; Pred. No. 5.3e-23;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37
Db 114 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 150

RESULT 10
US-08-464-023A-3
; Sequence 3, Application US/08464023A
; Patent No. 5854399
; GENERAL INFORMATION:
; APPLICANT: Salomon, David S.
```

CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,001  
FILING DATE: 23-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/154,198  
FILING DATE: 17-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 15280-63-1-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-464-023A-5

Query Match 100.0%; Score 235; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 5.3e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKKCSLCKWHGQRCFPPQAFLEPGCD 37  
|||||  
Db 114 NCGSVPHDTWLPKKCSLCKWHGQRCFPPQAFLEPGCD 150  
|||||

RESULT 12  
US-08-471-371-7  
Sequence 7, Application US/08471371  
Patent No. 5981215  
GENERAL INFORMATION:  
APPLICANT: Meisner, Paul S.  
APPLICANT: Coleman, Timothy A.  
TITLE OF INVENTION: HUMAN CRIPTIN GROWTH FACTOR  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
ADDRESSEE: STEWART & OLSTEIN  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: N.J.  
COUNTRY: U.S.A.  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471.371  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferraro, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 32580-455  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 994-1700  
TELEFAX: (201) 994-1744  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-471-371-7

Query Match 100.0%; Score 235; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 5.3e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKKCSLCKWHGQRCFPPQAFLEPGCD 37  
|||||  
Db 114 NCGSVPHDTWLPKKCSLCKWHGQRCFPPQAFLEPGCD 150  
|||||

RESULT 13  
US-09-949-016-7062  
Sequence 7062, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR FILING DATE: 60/241,755  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7062  
LENGTH: 192  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-7062

Query Match 100.0%; Score 235; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 5.4e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKKCSLCKWHGQRCFPPQAFLEPGCD 37  
|||||  
Db 118 NCGSVPHDTWLPKKCSLCKWHGQRCFPPQAFLEPGCD 154  
|||||

RESULT 14  
US-09-949-016-8959  
Sequence 8959, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR FILING DATE: 60/241,755  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8959  
LENGTH: 192  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-8959

Query Match 100.0%; Score 235; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 5.4e-23;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKKCSLCKWHGQRCFPPQAFLEPGCD 37  
|||||

Db 118 NCGSVPHDTWLPKCSLCKCWHGQLRCPQAFPLPGCD 154

RESULT 15

US-09-949-016-11113  
 ; Sequence 11113, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 11113  
 ; LENGTH: 192  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-09-949-016-11113

Query Match 100.0%; Score 235; DB 2; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-23;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKCSLCKCWHGQLRCPQAFPLPGCD 37  
 Db 118 NCGSVPHDTWLPKCSLCKCWHGQLRCPQAFPLPGCD 154

Search completed: February 3, 2006, 18:33:13  
 Job time : 10.044 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:15:30 ; Search time 35.6844 Seconds  
(without alignments)  
731.539 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_114\_150

Perfect score: 235

Sequence: 1 NCGSVPHDTWLPKKSCLKCWGQLRCFPQAFILPGCD 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	235	100.0	188	1	PI3385 homo sapien
2	235	100.0	188	1	TDGF1_HUMAN
3	235	100.0	188	2	PI51864 homo sapien
4	227	96.6	181	2	O8TCC1 homo sapien
5	194	82.6	171	1	Q58D57 BOVIN
6	194	82.6	171	1	TDGF1_MOUSE
7	168	71.5	66	2	Q7TQ06 mus musculus
8	123	52.3	183	2	Q91AT2 BRARE
9	123	52.3	183	2	O57517 BRACHYDANIO
10	123	52.3	183	2	O50415 BRARE
11	102	43.4	193	2	O57516 BRACHYDANIO
12	88.5	37.7	190	2	Q91803_CHICK
13	88.5	37.7	190	2	Q800J2_XENLA
14	88	37.4	223	2	Q91649_XENLA
15	88	37.4	223	2	O53T05_HUMAN
16	83	35.3	202	2	O9GZK3_HUMAN
17	82	34.9	475	2	P97766_MOUSE
18	82	34.9	4007	1	Q4W596_HUMAN
19	78	33.2	814	2	FRAS1_HUMAN
20	77	32.8	1311	2	Q6ZWJ8_HUMAN
21	77	32.8	1527	2	Q961K8_DROME
22	73.5	31.3	1535	2	Q9VZ24_DROME
23	71	30.2	4010	1	Q23991_DROME
24	70	29.8	1254	2	Q80T14 mus musculus
25	69	29.4	318	2	Q56N18_MOUSE
26	69	29.4	668	2	Q4SH80_TETNG
27	68	28.9	1431	2	Q5D734_BRARE
28	66	28.1	1027	1	Q80U60_MOUSE
29	65.5	27.9	350	2	Q7C342 BRACHYDANIO
30	65.5	27.9	1048	1	Q5FW56_MOUSE
31	65	27.7	218	2	CRIM1_CHICK
					Q96GF5_HUMAN

32	65	27.7	719	2	Q4S463_TETNG
33	65	27.7	1496	2	Q92626_HUMAN
34	64.5	27.4	281	2	Q9B290_HUMAN
35	64.5	27.4	308	2	O6WN33_HUMAN
36	64.5	27.4	429	1	CRDL2_HUMAN
37	64.5	27.4	915	2	Q4RSI5_TETNG
38	64	27.2	741	2	Q4RLP8_TETNG
39	63	26.8	620	2	Q4WDC1_ASPTU
40	63	26.8	1166	2	Q57UP5_9TRYP
41	63	26.8	1449	2	Q640B2_XENTR
42	62.5	26.6	245	2	Q7TPA0_RAT
43	62.5	26.6	640	2	Q96397_CHLRE
44	62.5	26.6	784	2	Q8BM43_MOUSE
45	62.5	26.6	814	2	Q7ZX15_XENLA

#### ALIGNMENTS

RESULT 1	ID	TDGF1_HUMAN	STANDARD;	PRT;	188 AA.
AC	PI3385;				
DT	01-JAN-1990	(Rel. 13, Created)			
DT	01-JAN-1990	(Rel. 13, Last sequence update)			
DT	10-MAY-2005	(Rel. 47, Last annotation update)			
DE	Teratocarcinoma-derived growth factor 1 precursor (Epidermal growth factor-like cripto protein CR1) (Cripto-1 growth factor) (CRGF).				
DE	Name=TDGF1; Synonyms=CR1PTO;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
OC	Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RX	MEDLINE=9005403; PubMed=2792079;				
RA	Ciccodicola A., Dono R., Obici S., Zollo M., Persico M.G.;				
RT	"Molecular characterization of a gene of the 'EGF family' expressed in undifferentiated human NTera2 teratocarcinoma cells.";				
RL	EMBO J. 8:1987-1991(1989).				
RN	[2]				
RP	NUCLEOTIDE SEQUENCE.				
RX	MEDLINE=9135371; PubMed=1882841;				
RA	Dono R., Montuori N., Rocchi M., de Ponti-Zilli L., Ciccodicola A., Persico M.G.;				
RT	"Isolation and characterization of the CR1PTO autosomal gene and its X-linked related sequence.";				
RL	Am. J. Hum. Genet. 49:555-565(1991).				
RN	[3]				
RP	PROTEIN SEQUENCE OF 31-45.				
RX	PubMed=15340161; DOI=10.1110/ps.04682504;				
RA	Zhang Z., Henzel W.J.;				
RT	"Signal peptide prediction based on analysis of experimentally verified cleavage sites.";				
RL	Protein Sci. 13:2819-2824(2004).				
RN	[4]				
RP	DISULFIDE BONDS.				
RX	MEDLINE=22800780; PubMed=12919325;				
RA	DOI=10.1046/j.1432-1033.2003.03749.x;				
RT	Foley S.F., Van Vlijmen H.W., Boynton R.E., Adkins H.B., Cheung A.E., Singh J., Sanicola M., Young C.N., Wen D.;				
RL	"The CR1PTO/FR1-1/CRYPTIC (CFC) domain of human Cripto.";				
RN	Eur. J. Biochem. 270:3610-3618(2003).				
CC	-1- FUNCTION: Could play a role in the determination of the epiblastic cells that subsequently give rise to the mesoderm.				
CC	-1- TISSUE SPECIFICITY: Preferentially expressed in gastric and colorectal carcinomas than in their normal counterparts.				
CC	-1- SIMILARITY: Contains 1 EGF-like domain.				
CC	-----				
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its				





DE Teratocarcinoma-derived growth factor 1.  
GN Name=TDGF1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Brain, and Kidney;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences.";  
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Kidney;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Brain;  
RG NIH MGC Project;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022393; AAH22393.1; -; mRNA.  
DR EMBL; BC067844; AAH67844.1; -; mRNA.  
DR PIR; A39787; A39787.  
DR HSP; P00740; IEDM.  
DR Ensembl; ENSG00000163828; Homo sapiens.  
DR InterPro; IPR00742; EGF\_2.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR006210; IEGF.  
DR Pfam; PF00008; EGF; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS00026; EGF\_3; 1.  
KW EGF-like domain; Transmembrane.  
SQ SEQUENCE 188 AA; 21141 MW; E1B3A8461C855FFF CRC64;  
Query Match 100.0%; Score 235; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 8.6e-22;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 NCGSVPHDTWLPKCSLCKWHGQLRCFPQAFPGCD 37  
|||||  
Db 114 NCGSVPHDTWLPKCSLCKWHGQLRCFPQAFPGCD 150  
|||||  
RESULT 4  
ID Q58D57\_BOVIN PRELIMINARY; PRT; 181 AA.  
AC Q58D57;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE Teratocarcinoma-derived growth factor 1.

GN Name=TDGF1;  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
OC Pecora; Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Pooled;  
RX MEDLINE=21180013; PubMed=11282978; DOI=10.1101/gr.170101;  
RA Smith T.P.L., Grosse W.M., Freking B.A., Roberts A.J., Stone R.T.,  
RA Casas E., Wray J.E., White J., Cho J., Fahrkrug S.C., Bennett G.L.,  
RA Heaton M.P., Laegreid W.W., Rohrer G.A., Chitko-McKown C.G.,  
RA Perteira G., Holt I., Karamycheva S., Liang F., Quackenbush J.,  
RA Keele J.W.;  
RT "Sequence evaluation of four pooled-tissue normalized bovine cDNA  
RL libraries and construction of a gene index for cattle.";  
RN Genome Res. 11:626-630 (2001).  
[2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Pooled;  
RA Harhay G.P., Sonstegard T.S., Van Tassel C.P., Clawson M.L.,  
RA Heaton M.P., Keele J.W., Snelling W.M., Weidmann R.T., Smith T.P.L.;  
RT "Sequencing and analysis of Bos taurus full-length insert cDNA  
RL clones.";  
RN Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BT021740; AAX46587.1; -; mRNA.  
SQ SEQUENCE 181 AA; 20023 MW; A4E8AEA78E106514 CRC64;  
Query Match 96.6%; Score 227; DB 2; Length 181;  
Best Local Similarity 91.9%; Pred. No. 8.5e-21;  
Matches 34; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 NCGSVPHDTWLPKCSLCKWHGQLRCFPQAFPGCD 37  
|||||  
Db 109 NCGSVPHDTWLPKCSLCKWHGQLRCFPQAFPGCD 145  
|||||  
RESULT 5  
ID TDGF1\_MOUSE STANDARD; PRT; 171 AA.  
AC P51865;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Teratocarcinoma-derived growth factor precursor (Epidermal growth  
DE factor-like Cripito protein) (Cripito growth factor).  
GN Name=tdgf1; Synonyms=Cripito;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muroidae; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6; TISSUE=Embryo;  
RX MEDLINE=94094736; PubMed=7916676;  
RA Dono R., Scalera L., Pacifico F., Acampora D., Persico M.G.,  
RA Simeone A.;  
RT "The murine cripito gene: expression during mesoderm induction and  
RT early heart morphogenesis.";  
RL Development 118:1157-1168 (1993).  
RN [2]  
RP NUCLEOTIDE SEQUENCE OF 135-171.  
RC STRAIN=129/SV;  
RX MEDLINE=96285959; PubMed=8661720; DOI=10.1007/s003359900100;  
RA Liguori G., Tucci M., Montuori N., Dono R., Lago C.I., Pacifico A.F.,  
RA Persico M.G.;  
RT "Characterization of the mouse Tdgfl gene and Tdgf pseudogenes.";  
RN Mamm. Genome 7:344-348 (1996).  
CC -1- FUNCTION: Could play a role in the determination of the epiblastic  
CC cells that subsequently give rise to the mesoderm.  
CC -1- TISSUE SPECIFICITY: Expressed at low level in specific organs of

the adult animal such as spleen, heart, lung and brain. During gastrulation, expressed in the forming mesoderm. In later stages of the developing heart, expression is restricted to the truncus arteriosus.

-!- DEVELOPMENTAL STAGE: First expressed prior to the onset of gastrulation (early streak stage), then continues throughout embryonic development.

-!- SIMILARITY: Contains 1 EGF-like domain.

-----

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

-----

EMBL; M87321; AAA37459.1; -; mRNA.  
EMBL; X94083; CAA63827.1; -; Genomic\_DNA.  
PIR; I49612; I49612.  
HSP; P00740; 1EDM.  
Ensembl; ENSMUSG0000032494; Mus musculus.  
MGI; MGI:98658; Tdgfl.  
GO; GO:0005615; C:extracellular space; TAS.  
GO; GO:0019897; C:extrinsic to plasma membrane; IDA.  
GO; GO:0001763; P:branching morphogenesis; IDA.  
GO; GO:0030154; P:cell differentiation; TAS.  
GO; GO:0008595; P:determination of anterior/posterior axis, e. . .; IMP.  
GO; GO:0007507; P:heart development; IDA.  
GO; GO:0030335; P:positive regulation of cell migration; IDA.  
InterPro; IPR000742; EGF 2.  
InterPro; IPR006209; EGF\_like.  
Pfam; PF00008; EGF; 1.  
PROSITE; PS00022; EGF\_1; 1.  
PROSITE; PS50026; EGF\_3; 1.  
EGF-like domain; Glycoprotein; Growth factor; Signal.  
SIGNAL 1 17 Potential.  
CHAIN 18 171 Teratocarcinoma-derived growth factor.  
DOMAIN 62 91 EGF-like.  
CARBOHYD 63 63 N-linked (GlcNAc . .) (Potential).  
DISULFID 66 73 By similarity.  
DISULFID 67 79 By similarity.  
DISULFID 81 90 By similarity.  
DISULFID 99 117 By similarity.  
DISULFID 112 133 By similarity.  
DISULFID 115 124 By similarity.  
SEQUENCE 171 AA; 18754 MW; C52051AEACDB5380 CRC64;

Query Match 82.6%; Score 194; DB 1; Length 171;  
Best Local Similarity 78.4%; Pred. No. 1.2e-16;  
Matches 29; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NCGSPVHDTWLPKCSLCKCHWGQLRCPQAFPLPGCD 37  
:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 98 HCGSILHGTWLPKCSLCKCHWGQLRCPQAFPLPGCD 134

RESULT 6  
Q7TQ06\_MOUSE PRELIMINARY; PRT; 171 AA.  
AC Q7TQ06;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Tdgfl protein.  
GN Name=Tdgfl;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=blastocyst;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.B., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=blastocyst;  
RA Strausberg R.;  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC052646; AAH52646.1; -; mRNA.  
DR HSP; P00740; 1EDM.  
DR MGI; MGI:98658; Tdgfl.  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR GO; GO:0019897; C:extrinsic to plasma membrane; IDA.  
DR GO; GO:0008595; P:determination of anterior/posterior axis, e. . .; IMP.  
DR InterPro; IPR000742; EGF 2.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR006210; EGF.  
DR Pfam; PF00008; EGF; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS50026; EGF\_3; 1.  
SEQUENCE 171 AA; 18654 MW; C53400EBACDB5380 CRC64;

Query Match 82.6%; Score 194; DB 2; Length 171;  
Best Local Similarity 78.4%; Pred. No. 1.2e-16;  
Matches 29; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NCGSPVHDTWLPKCSLCKCHWGQLRCPQAFPLPGCD 37  
:||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 98 HCGSILHGTWLPKCSLCKCHWGQLRCPQAFPLPGCD 134

RESULT 7  
Q5IAT2\_BRARE PRELIMINARY; PRT; 66 AA.  
AC Q5IAT2;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Crpto (Fragment).  
GN Name=crpto;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Del Giacco L., Cotelli P.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF151352; AAF73188.1; -; mRNA.  
DR ZFIN; ZDB-GENE-000710-5; crpto.  
FT NON\_TER 1 1  
FT NON\_TER 66 66  
SEQUENCE 66 AA; 7275 MW; 5C6ABB16D5832871 CRC64;



SQ SEQUENCE 190 AA; 21337 MW; 6079CASFBCIF0A83 CRC64;  
 Query Match 37.7%; Score 88.5; DB 2; Length 190;  
 Best Local Similarity 42.1%; Pred. No. 0.0033;  
 Matches 16; Conservative 6; Mismatches 13; Indels 3; Gaps 2;

```

QY      1 NCGSPVPHDTWLPKKCSLCKCWGQLRCF-PQAFLPGCD 37
          : ||| | : : ||| : ||| | : : |||
DB     115 SCSGVPHGDWIRQGCLLCRCVSGVLHCFKPES--EDCD 150

```

RESULT 13  
 Q91649\_XENLA PRELIMINARY; PRT; 190 AA.  
 ID Q91649\_XENLA PRELIMINARY;  
 AC Q91649;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Fibroblast growth factor receptor ligand 1.  
 GN Name=FRU1;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=96069863; PubMed=7585965; DOI=10.1016/0092-8674(95)90102-7;  
 RA Kinoshita N., Minshull J., Kirchner M.W.;  
 RT "The identification of two novel ligands of the FGF receptor by a  
 RL yeast screening method and their activity in xenopus development."; Cell 83:621-630 (1995).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RX PubMed=11024280; DOI=10.1016/S0378-1119(00)00337-1;  
 RA Colas J.-F., Schoenwolf G.C.;  
 RT "Subtractive hybridization identifies chick-cripto, a novel EGF-CFC  
 RL ortholog expressed during gastrulation, neurulation and early  
 RN cardiogenesis."; Gene 255:205-217 (2000).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Colas J.-F., Schoenwolf G.C.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U40457; AAA83569.1; -; mRNA.  
 DR EMBL; AF234786; AAF78013.1; -; Genomic\_DNA.  
 DR HSSP; P00749; IURK.  
 DR GO; GO:0004872; Fireceptor activity; IEA.  
 DR InterPro; IPRO06209; EGF\_like.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 KW Receptor.  
 SQ SEQUENCE 190 AA; 21462 MW; 6FC1660D2B953176 CRC64;
 

Query Match 37.7%; Score 88.5; DB 2; Length 190;  
 Best Local Similarity 42.1%; Pred. No. 0.0033;  
 Matches 16; Conservative 6; Mismatches 13; Indels 3; Gaps 2;

```

QY      1 NCGSPVPHDTWLPKKCSLCKCWGQLRCF-PQAFLPGCD 37
          : ||| | : : ||| : ||| | : : |||
DB     115 SCSGVPHGDWIRQGCLLCRCVSGVLHCFKPES--EDCD 150

```

RESULT 14  
 Q53T05\_HUMAN PRELIMINARY; PRT; 223 AA.  
 ID Q53T05\_HUMAN PRELIMINARY;  
 AC Q53T05;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Hypothetical protein CRYPTIC.  
 GN Name=CRIPTIC;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OX Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RA Cotton M., Du H.;  
RT "The sequence of Homo sapiens BAC clone RP11-356110.";  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN (2)  
RP NUCLEOTIDE SEQUENCE.  
RA Waterston R.H.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
RN (3)  
RP NUCLEOTIDE SEQUENCE.  
RA Waterston R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
RN (4)  
RP NUCLEOTIDE SEQUENCE.  
RA Wilson R.K.;  
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC013269; AAY14955.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 223 AA; 24642 MW; 457EB443BC0CF1A5 CRC64;  
  
Query Match 37.4%; Score 88; DB 2; Length 223;  
Best Local Similarity 38.9%; Pred. No. 0.0045;  
Matches 14; Conservative 4; Mismatches 18; Indels 0; Gaps 0;  
  
Qy 2 CGSVPHDTPKCKSLCKCKWHGQLRCPPQAFPLPGCD 37  
Db 123 CGALEHGAWTIRACHLCRCIFGALHCLPLQTPDRCD 158  
  
RESULT 15  
O9GZ33 HUMAN  
ID O9GZ33 HUMAN PRELIMINARY; PRT; 223 AA.  
AC O9GZ33;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Cryptic.  
GN Name=CFC1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20517351; PubMed=11062482; DOI=10.1038/81695;  
RA Bamford R.N., Roessler E., Burdine R., Saplakoglu U., Cruz J.D.,  
RA Splitt M., Towbin J., Marino B., Schier A.F., Shen M.M., Muenke M.,  
RA Casey B.;  
RT "Loss-of-function mutations in the EGF-CFC gene CFC1 are associated  
RT with human left-right laterality defects.";  
RL Nat. Genet. 26:365-369(2000).  
RN (2)  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Lung, and PCR rescued clones;  
RX MEDLINE=22388257; PubMed=12477933; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN (3)  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=PCR rescued clones;  
RA Director MGC Project;  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
RN (4)  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Lung;  
RA Director MGC Project;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
RN (5)  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Lung;  
RG NIH MGC Project;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF312925; AAG42475.1; -; Genomic\_DNA.  
DR EMBL; AF312769; AAG30294.1; -; mRNA.  
DR EMBL; BC069508; AAH69508.1; -; mRNA.  
DR EMBL; BC074825; AAH74825.1; -; mRNA.  
DR EMBL; BC074826; AAH74826.1; -; mRNA.  
DR HGSP; P00750; 1TPG.  
DR HGNC; HGNC:18292; CFC1.  
DR GO; GO:0007368; P:determination of left/right symmetry; NAS.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR006210; IEGF.  
DR Pfam; PF00008; EGF\_1.  
DR SMART; SM00181; EGF\_1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS00026; EGF\_3; 1.  
KW EGF-like domain; Transmembrane.  
SQ SEQUENCE 223 AA; 24612 MW; B52852A00ABCFLA3 CRC64;  
  
Query Match 37.4%; Score 88; DB 2; Length 223;  
Best Local Similarity 38.9%; Pred. No. 0.0045;  
Matches 14; Conservative 4; Mismatches 18; Indels 0; Gaps 0;  
  
Qy 2 CGSVPHDTPKCKSLCKCKWHGQLRCPPQAFPLPGCD 37  
Db 123 CGALEHGAWTIRACHLCRCIFGALHCLPLQTPDRCD 158

Search completed: February 3, 2006, 18:29:59

Job time : 37.6844 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:23:01 ; Search time 6.08444 Seconds  
(without alignments)  
585.102 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_114\_150

Perfect score: 235

Sequence: 1 NCGSVPHDTWLPKCSLCKCMHGLRCFPOAFLPGCD 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	235	100.0	188	2 A39787	teratocarcinoma-de
2	235	100.0	188	2 A30362	teratocarcinoma-de
3	194	82.6	171	2 T49612	teratocarcinoma-de
4	73.5	31.3	1535	2 S46224	peroxidase - fru1
5	62.5	26.6	640	2 T08179	LRG5 protein - chl
6	60	25.5	1464	1 CGHUIS	collagen alpha 1(I
7	59.5	25.3	835	2 JP0076	nel protein - chic
8	59	25.1	153	2 A27179	collagen alpha 1(I
9	58.5	24.9	847	2 A56039	GTPase-activating
10	58	24.7	98	2 T49562	alpha-1 type III c
11	58	24.7	488	2 A27353	collagen alpha 1(I
12	58	24.7	615	2 A05269	collagen alpha 1(I
13	58	24.7	886	2 T50694	collagen alpha 1(I
14	58	24.7	1464	2 S59856	collagen alpha 1(I
15	58	24.7	1466	1 CGHUTL	collagen alpha 1(I
16	58	24.7	1497	2 T49607	procollagen type V
17	57	24.3	312	2 T35887	hypothetical prote
18	57	24.3	837	2 T19271	hypothetical prote
19	57	24.3	1496	1 CGHUTV	collagen alpha 2(V
20	56.5	24.0	381	2 E83397	pyrroloquinoline q
21	56.5	24.0	445	2 A56043	steroid hormone re
22	56.5	24.0	1453	2 S21626	collagen alpha 1(I
23	56	23.8	161	2 T16847	hypothetical prote
24	56	23.8	1486	1 B40333	collagen alpha 1(I
25	56	23.8	1487	2 B41182	collagen alpha 1(I
26	56	23.8	1492	2 A40333	collagen alpha 1(I
27	55.5	23.6	260	2 C86228	hypothetical prote
28	55.5	23.6	505	2 C64502	hypothetical prote
29	55.5	23.6	747	2 T16274	hypothetical prote

30 55.5 23.6 1449 2 B84426 hypothetical prote  
31 55.5 23.6 1816 1 B89960 laminin alpha-4 ch  
32 55.5 23.6 2205 1 MNWVRN nonstructural poly  
33 55.5 23.6 5376 2 T42215 zonadhesin - mouse  
34 55 23.4 105 2 T47210 ccg-4 protein 1 li  
35 55 23.4 181 2 B84567 probable RING zinc  
36 55 23.4 206 2 S18250 collagen alpha 1(I  
37 55 23.4 810 2 T10756 Nel-homolog protei  
38 55 23.4 926 2 T39082 hypothetical prote  
39 55 23.4 1352 2 G84473 hypothetical prote  
40 55 23.4 1487 1 CGHU6C collagen alpha 1(I  
41 54 23.0 333 2 AG2214 hypothetical prote  
42 54 23.0 426 2 T48379 gene hb protein -  
43 54 23.0 2318 2 S45306 notch3 protein - h  
44 54 23.0 2321 2 S78549 conserved hypothet  
45 53.5 22.8 161 2 A82864

## ALIGNMENTS

## RESULT 1

A39787

teratocarcinoma-derived growth factor - human

C;Species: Homo sapiens (man)

C;Date: 20-Mar-1992 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004

C;Accession: A39787

R;Dono, R.; Montuori, N.; Rocchi, M.; De Ponti-Zilli, L.; Ciccodicola, A.; Persico, M.G.

A;Title: Isolation and characterization of the CRIPTO autosomal gene and its X-linked re

A;Reference number: A39787; MUID:91353571; PMID:1882841

A;Accession: A39787

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-188 &lt;DON&gt;

A;Cross-references: UNIPROT:Q8TCCI; UNIPARC:UPI00000145185

A;Note: the authors translated the codon GAC for residue 43 as Glu

C;Superfamily: teratocarcinoma-derived growth factor 1; EGF homology

Query Match 100.0%; Score 235; DB 2; Length 188;

Best Local Similarity 100.0%; Pred. No. 4.1e-21;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKCSLCKCMHGLRCFPOAFLPGCD 37

Db 114 NCGSVPHDTWLPKCSLCKCMHGLRCFPOAFLPGCD 150

## RESULT 2

A30362

teratocarcinoma-derived growth factor 1 - human

N;Alternate names: CRIPTO protein

C;Species: Homo sapiens (man)

C;Date: 18-Oct-1989 #sequence\_revision 18-Oct-1989 #text\_change 09-Jul-2004

C;Accession: A39787; A30362

R;Dono, R.; Montuori, N.; Rocchi, M.; De Ponti-Zilli, L.; Ciccodicola, A.; Persico, M.G.

A;Title: Isolation and characterization of the CRIPTO autosomal gene and its X-linked re

A;Reference number: A39787; MUID:91353571; PMID:1882841

A;Accession: B39787

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-188 &lt;DON&gt;

A;Cross-references: UNIPROT:P13385; UNIPARC:UPI000004966D; GB:M96955; GB:M37099; NID:g33

R;Ciccodicola, A.; Dono, R.; Obici, S.; Simeone, A.; Zollo, M.; Persico, M.G.

EMBO J. 8, 1987-1991, 1989

A;Title: Molecular characterization of a gene of the 'EGF family' expressed in undifferen

A;Reference number: A30362; MUID:90005403; PMID:2792079

A;Accession: A30362

A;Molecule type: mRNA

A;Residues: 1-188 &lt;CIC&gt;

A;Cross-references: UNIPARC:UPI000004966D; GB:X14253; NID:g30220; PIDN:CAA32457.1; PID:g

C;Superfamily: teratocarcinoma-derived growth factor 1; EGF homology





R;Chu, M.L.; de Wet, W.; Bernard, M.; Ramirez, F.  
J. Biol. Chem. 260, 2315-2320, 1985  
A:Title: Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promoter  
A:Reference number: 155237; MUID:85130970; PMID:2857713  
A:Accession: 155237  
A:Status: translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-34 <CH2>  
A:Cross-references: UNIPARC:UPI000016A6B1; GB:M10627; NID:gl180383; PIDN:AAAS1992.1; PID:  
R;Wirtz, M.K.; Keene, D.R.; Hori, H.; Glanville, R.W.; Steinmann, B.; Rao, V.H.; Hollis  
J. Biol. Chem. 265, 6312-6317, 1990  
A:Title: In vivo and in vitro noncovalent association of excised alpha(I) amino-termina  
rome, type VII.  
A:Reference number: A35233; MUID:90202908; PMID:2318855  
A:Accession: A35233  
A:Molecule type: protein  
A:Residues: 33-52 <WIR>  
A:Cross-references: UNIPARC:UPI0000173B3D  
A:Note: this propeptide fragment remained non-covalently bound to a defective, uncleaved  
R;Well, D.; d'Alessio, M.; Ramirez, F.; de Wet, W.; Cole, W.G.; Chan, D.; Bateman, J.F.  
EMBO J. 8, 1705-1710, 1989  
A:Title: A base substitution in the exon of a collagen gene causes alternative splicing  
A:Reference number: S09400; MUID:89356643; PMID:2767050  
A:Accession: S09400  
A:Molecule type: mRNA  
A:Residues: 156-183 <WEI>  
A:Cross-references: UNIPARC:UPI0000173B3E  
R;Click, E.M.; Bornstein, P.  
Biochemistry 9, 4599-4706, 1970  
A:Title: Isolation and characterization of the cyanogen bromide peptides from the alpha  
A:Reference number: A90567; MUID:17038625; PMID:5529814  
A:Contents: CNBr0-1, CNBr2, CNBr4, CNBr5  
A:Accession: B90567  
A:Molecule type: protein  
A:Residues: 162-198, 'Z', 200-201, 'Z', 203-206, 'Z', 208-209, 'Z', 211-228, 'B', 230, 'BB', 233, 'Z',  
A:Cross-references: UNIPARC:UPI0000173B3F  
A:Experimental source: skin  
R;Baetge, B.; Notbohm, H.; Diebold, J.; Lehmann, H.; Bodo, M.; Deutzmann, R.; Mueller, E.  
Eur. J. Biochem. 192, 153-159, 1990  
A:Title: A critical crosslink region in human-bone-derived collagen type I. Specific cle  
A:Reference number: S11372; MUID:90382436; PMID:2169412  
A:Accession: S11372  
A:Molecule type: protein  
A:Residues: 175-187; 274-287, 'P', 289 <BAE>  
A:Cross-references: UNIPARC:UPI0000173B40; UNIPARC:UPI0000173B41  
A:Note: sequence of collagen alpha 1(S)(I) isolated from bone after pepsin digestion  
R;Deak, S.B.; Scholz, P.M.; Amenta, P.S.; Constantinou, C.D.; Levi-Minzi, S.A.; Gonzalez  
J. Biol. Chem. 266, 21827-21832, 1991  
A:Title: The substitution of arginine for glycine 85 of the alpha 1(I) procollagen chain  
operative melting of intact type I collagen.  
A:Reference number: 155342; MUID:92042092; PMID:1718984  
A:Accession: 155342  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 258-268; 1347-1357 <DEA>  
A:Cross-references: UNIPARC:UPI0000173B42; UNIPARC:UPI0000173B43; GB:S67495; NID:9239007  
A:Note: sequences from the 5' and 3' ends only are shown; mutant sequence 263-Arg report  
R;Morgan, P.H.; Jacobs, H.G.; Segrest, J.P.; Cunningham, L.W.  
J. Biol. Chem. 245, 5042-5048, 1970  
A:Title: Comparative study of glycopeptides derived from selected vertebrate collagens.  
A:Reference number: A92069; MUID:71001508; PMID:4319110  
A:Accession: A92069  
A:Molecule type: protein  
A:Residues: 263-268 <MOR>  
A:Cross-references: UNIPARC:UPI000014DF11  
A:Experimental source: skin  
A:Note: attachment of 2-O-alpha-D-glucosyl-O-beta-D-galactose to 5-hydroxylysine  
R;Labhard, M.E.; Hollister, D.W.  
Matrix 10, 124-130, 1990  
A:Title: Segmental amplification of the entire helical and telopeptide regions of the ch  
A:Reference number: S15989; MUID:90326017; PMID:2374517  
A:Accession: S15989

A:Molecule type: mRNA  
A:Residues: 281-302,402-420,823-843,925-944,1026-1045,1143-1162 <LAB>  
A:Cross-references: UNIPARC:UPI0000173B44; UNIPARC:UPI0000173B45; UNIPARC:UPI0000173B46;  
R:Wirtz, M.K.; Rao, V.H.; Rao, V.H.; Llabhard, R.W.; Labhard, M.E.; Pretorius, P.J.; de Vries, W.N.  
Connect. Tissue Res. 29, 1-11, 1993  
A:Title: A cysteine for glycine substitution at position 175 in an alpha 1 (I) chain of  
A:Reference number: 152905; MUID:93339042; PMID:8339541  
A:Accession: 152905  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 342-352, 'C', 354-359 <W12>  
A:Cross-references: UNIPARC:UPI00000708E6; GB:S64717; NID:g408195; PIDN:AAB27677.1; PID:  
A:Note: mutant sequence from patient with osteogenesis imperfecta  
E:Bernard, M.P.; Chu, M.L.; Myers, J.C.; Ramirez, F.; Eikenberry, E.F.; Prockop, D.J.  
Biochemistry 22, 5213-5223, 1983  
A:Title: Nucleotide sequences of complementary deoxyribonucleic acids for the proalpha1  
A:Reference number: A90476; MUID:84080385; PMID:6689127  
A:Accession: A90476  
A:Molecule type: mRNA  
A:Residues: 425-1250, 'X', 1252-1328, 'S', 1330-1390, 'X', 1392-1464 <BER>  
A:Cross-references: UNIPARC:UPI0000173B4A; GB:K01228; NID:g180391; PIDN:AAAS1995.1; PID:  
A:Note: sequence partially completed for missing nucleotides by A29439  
R:Chu, M.L.; Gargiulo, V.; Williams, C.J.; Ramirez, F.  
J. Biol. Chem. 260, 691-694, 1985  
A:Title: Multixon deletion in an osteogenesis imperfecta variant with increased type II  
A:Reference number: A22161; MUID:85104934; PMID:2981843  
A:Accession: A22161  
A:Molecule type: DNA  
A:Residues: 472-594, 'R', 596-607 <CH3>  
A:Cross-references: UNIPARC:UPI000011F796; GB:K03178; GB:K03179; NID:g179612; NID:g17961  
A:Note: the authors translated the codon CGT for residue 595 as Pro  
R:Wallis, G.A.; Starman, B.J.; Zinn, A.B.; Byers, P.H.  
Am. J. Hum. Genet. 46, 1034-1040, 1990  
A:Title: Variable expression of osteogenesis imperfecta in a nuclear family is explained  
A:Reference number: A35336; MUID:90252792; PMID:2339700  
A:Accession: A35336  
A:Molecule type: mRNA  
A:Residues: 710-720, 'R', 722-737, 'E', 739-745 <WAL>  
A:Cross-references: UNIPARC:UPI0000173B4B  
A:Note: the authors translated the codons CAG for 721 and CGT for 738 as Glu  
R:Forlino, A.; Zolezzi, F.; Valli, M.; Pignatti, P.F.; Cetta, G.; Brunelli, P.C.; Mottes  
Hum. Mol. Genet. 3, 2201-2206, 1994  
A:Title: Severe (type III) osteogenesis imperfecta due to glycine substitutions in the c  
A:Reference number: I54365; MUID:95187161; PMID:7881420  
A:Accession: I54365  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 746-766, 'S', 768-781 <FOR>  
A:Cross-references: UNIPARC:UPI000016A6FA; GB:L47667; NID:g1009093; PIDN:AAB59576.1; PID:  
R:Chesler, S.D.; Wallis, G.A.; Byers, P.H.  
J. Biol. Chem. 268, 18218-18225, 1993  
A:Title: Mutations in the carboxyl-terminal propeptide of the pro alpha 1(I) chain of ty  
A:Reference number: A47426; MUID:93352846; PMID:8349697  
A:Accession: A47426  
A:Molecule type: mRNA  
A:Residues: 1179-1276, 'H', 1278-1336, 1339-1387, 'R', 1389-1464 <CHS>  
A:Cross-references: UNIPARC:UPI0000073A2A; GB:S64596; NID:g407589; PIDN:AAB27856.1; PID:  
A:Note: sequence extracted from NCBI backbone (NCBI:136444, NCBI:136445)  
A:Note: does not represent an experimentally determined sequence but three different mut  
A:Accession: B47426  
A:Molecule type: mRNA  
A:Residues: 1179-1464 <CHA>  
A:Cross-references: UNIPARC:UPI0000173B4C  
A:Experimental source: normal dermal fibroblast culture  
A:Accession: C47426  
A:Molecule type: mRNA  
A:Residues: 1179-1276, 'H', 1278-1464 <CHS>  
A:Cross-references: UNIPARC:UPI0000173B4D  
A:Experimental source: fetal cell 86-237  
A:Accession: D47426  
A:Molecule type: mRNA  
A:Residues: 1179-1336, 1339-1464 <CHG>

A;Cross-references: UNIPARC:UPI0000173B4E  
A;Experimental source: fetal cell 86-146  
A;Accession: E47426  
A;Molecule type: mRNA  
A;Residues: 1179-1387, 'R', 1389-1464 <CH7>  
A;Cross-references: UNIPARC:UPI0000173B4F  
A;Experimental source: fetal cell 88-251  
R;Cohn, D.H.; Apone, S.; Eyre, D.R.; Starman, B.J.; Andreassen, P.; Charbonneau, H.; Nic  
J. Biol. Chem. 263, 14605-14607, 1988

Query Match 25.5%; Score 60; DB 1; Length 1464;  
Best Local Similarity 40.0%; Pred. No. 22;  
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 8 DTWLPKCSLCKCWHLQRLC 27  
DB 51 DVMKPEPCRICVDNGKVL 70

RESULT 7  
JF0076  
nei protein - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 24-Feb-1995 #sequence\_revision 24-Feb-1995 #text\_change 04-Apr-2004  
C;Accession: A38963; JP0076  
R;Matsumashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.; Taniguchi, S.; Hori, K.  
Dev. Dyn. 203, 212-222, 1995  
A;Title: New gene, nel, encoding a Mr 93K protein with EGF-like repeats is strongly exp  
A;Reference number: A38963; MUID:95383734; PMID:7655083  
A;Accession: A38963  
A;Molecule type: mRNA  
A;Residues: 1-835 <MAT>  
A;Cross-references: UNIPARC:UPI000017C023; DDBJ:D45365  
A;Experimental source: 9-day embryo  
R;Matsumashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.; Taniguchi, S.; Hori, K.  
submitted to JPIID, January 1995  
A;Description: A new gene, nel, encoding a Mr 93K protein with EGF-like repeats is stron  
A;Reference number: JP0076  
A;Accession: JP0076  
A;Molecule type: mRNA  
A;Residues: 1-835 <MA2>  
A;Cross-references: UNIPARC:UPI000017C023; DDBJ:D45365  
F;273-333/Domain: von Willebrand factor type C repeat homology <VWC>  
F;395-592/Region: EGF-like repeats  
F;444-480/Domain: EGF homology <EGFI>  
F;486-521/Domain: EGF homology <EGF>  
F;525-552/Domain: EGF homology <EGF2>

Query Match 25.3%; Score 59.5; DB 2; Length 835;  
Best Local Similarity 36.4%; Pred. No. 16;  
Matches 8; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 8 DTWLPKCSLCKCWHLQRLC 29  
DB 713 DSWI-QNCCQCRCLQGEVDCWP 733

RESULT 8  
A27179  
collagen alpha 1(I) chain precursor - chicken (fragment)  
C;Species: Gallus gallus (chicken)  
C;Date: 05-Oct-1988 #sequence\_revision 05-Oct-1988 #text\_change 09-Jul-2004  
C;Accession: A27179; A29367  
R;Finer, M.H.; Boedtker, H.; Doty, P.  
Gene 56, 71-78, 1987  
A;Title: Construction and characterization of cDNA clones encoding the 5' end of the chi  
A;Reference number: A27179; MUID:88056316; PMID:3678834  
A;Accession: A27179  
A;Molecule type: mRNA  
A;Residues: 1-153 <FIN>  
A;Cross-references: UNIPROT:P02457; UNIPARC:UPI00001712C5; GB:M17607; NID:G211473; PIDN:  
R;Finer, M.H.; Aho, S.; Gerstenfeld, L.C.; Boedtker, H.; Doty, P.

J. Biol. Chem. 262, 13323-13332, 1987  
A;Title: Unusual DNA sequences located within the promoter region and the first intron o  
A;Reference number: A29367; MUID:88007542; PMID:2820966  
A;Accession: A29367  
A;Molecule type: DNA  
A;Residues: 1-144 <FI2>  
A;Cross-references: UNIPARC:UPI0000177384; GB:M17607  
C;Genetics:  
A;Introns: 28/1; 93/1; 104/3; 116/3  
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproli  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-153/Domain: amino-terminal propeptide #status predicted <PRO>  
F;32-91/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match 25.1%; Score 59; DB 2; Length 153;  
Best Local Similarity 40.0%; Pred. No. 44;  
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 8 DTWLPKCSLCKCWHLQRLC 27  
DB 44 DVMKPEPCRICVDSGNILC 63

RESULT 9  
A56039  
GTPase-activating protein Gap1(m) - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 01-Dec-1995 #sequence\_revision 01-Dec-1995 #text\_change 09-Jul-2004  
C;Accession: A56039  
R;Maekawa, M.; Li, S.; Iwamatsu, A.; Morishita, T.; Yokota, K.; Imai, Y.; Kohsaka, S.; I  
Mol. Cell. Biol. 14, 6879-6885, 1994  
A;Title: A novel mammalian Ras GTPase-activating protein which has phospholipid-binding  
A;Reference number: A56039; MUID:95021216; PMID:7935405  
A;Accession: A56039  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-847 <MA5>  
A;Cross-references: UNIPROT:Q63713; UNIPARC:UPI00001351P6; GB:D30734; NID:G559374; PIDN:  
A;Experimental source: brain  
F;355-567/Domain: ras-specific GAP catalytic domain homology <GAP>  
F;602-702/Domain: pleckstrin repeat homology <PLK>

Query Match 24.9%; Score 58.5; DB 2; Length 847;  
Best Local Similarity 29.4%; Pred. No. 21;  
Matches 15; Conservative 6; Mismatches 13; Indels 17; Gaps 3;

QY 1 NCGSVPHDTWLPKCSLCKCWHLQRLC 36  
DB 688 NC--VEANEWIDMLCRVSRCHNRLSSPHPSAYLNGNWLCCQETSEGTGPGC 736

RESULT 10  
I49562  
alpha-1 type III collagen prepeptide - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I49562  
R;Lilau, G.; Yamada, Y.; de Crombrughe, B.  
J. Biol. Chem. 260, 531-536, 1985  
A;Title: Coordinate regulation of the levels of type III and type I collagen mRNA in mo  
A;Reference number: I49562; MUID:85080131; PMID:2981217  
A;Accession: I49562  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-98 <RES>  
A;Cross-references: UNIPROT:Q61369; UNIPARC:UPI0000028424; GB:K03038; NID:G192273; PIDN:  
C;Genetics:  
A;Introns: 29/1  
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
F;32-92/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match 24.7%; Score 58; DB 2; Length 98;

Best Local Similarity 40.0%; Pred. No. 3.9;  
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 8 DTWLPKCKSLCKWHGQLRC 27  
Db 44 DWRPEPCQICVDSGSVLC 63

RESULT 11  
A27353  
collagen alpha 1(III) chain precursor - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 28-Aug-1989 #sequence\_revision 30-Sep-1990 #text\_change 09-Jul-2004  
C;Accession: A27353; A22287; F50066  
R;Wood, L.; Theriault, N.; Vogeli, G.  
Gene 61, 225-230, 1987  
A;Title: Complete nucleotide sequence of the N-terminal domains of the murine alpha-1 ty  
A;Reference number: A91588; MUID:88167858; PMID:3443309  
A;Accession: A27353  
A;Molecule type: mRNA  
A;Residues: 1-488 <WOO>  
A;Cross-references: UNIPROT:P08121; UNIPARC:UPI000016CC49; GB:M18933; NID:G192276; PIDN:  
R;Liau, G.; Mudryj, M.; de Crombrughe, B.  
J. Biol. Chem. 260, 3773-3777, 1985  
A;Title: Identification of the promoter and first exon of the mouse alpha 1(III) collage  
A;Reference number: A92513; MUID:85131189; PMID:3972847  
A;Accession: A22287  
A;Molecule type: DNA  
A;Residues: 1-28 <LIA>  
A;Cross-references: UNIPARC:UPI0000177380  
C;Genetics:  
A;Introns: 29/1  
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
C;Keywords: coiled coil; extracellular matrix; glycoprotein; sulfoprotein; trimer; trip  
F;1-24/Domain: signal sequence #status predicted <SIG>  
F;25-154/Domain: amino-terminal propeptide #status predicted <APR>  
F;32-92/Domain: von Willebrand factor type C repeat homology <WVC>  
F;155-488/Product: collagen alpha 1(III) chain (fragment) #status predicted <MAT>  
F;164-488/Region: helical (fragment)  
F;40,152,159/Binding site: sulfate (Tyr) (covalent) #status predicted

Query Match 24.7%; Score 58; DB 2; Length 488;  
Best Local Similarity 40.0%; Pred. No. 15;  
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 8 DTWLPKCKSLCKWHGQLRC 27  
Db 44 DWRPEPCQICVDSGSVLC 63

RESULT 12  
A05269  
collagen alpha 1(III) chain precursor - chicken (fragments)  
C;Species: Gallus gallus (chicken)  
C;Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 09-Jul-2004  
C;Accession: A05269; A38035; A20855  
R;Yamada, Y.; Liau, G.; Mudryj, M.; Obici, S.; de Crombrughe, B.  
Nature 310, 333-337, 1984  
A;Title: Conservation of the sizes for one but not another class of exons in two chick  
A;Reference number: A05269; MUID:84270696; PMID:6547770  
A;Accession: A05269  
A;Molecule type: DNA  
A;Residues: 1-329 <YAM1>  
A;Cross-references: UNIPROT:P12105; UNIPARC:UPI0000177387  
A;Note: the authors translated the codon AAA for residue 68 as Arg, AAC for residue 173  
R;Yamada, Y.; Kuelen, K.; de Crombrughe, B.  
Nucleic Acids Res. 11, 2733-2744, 1983  
A;Reference number: A38035; MUID:83220816; PMID:6856474  
A;Accession: A38035  
A;Molecule type: DNA  
A;Residues: 330-615 <YAM2>  
A;Cross-references: UNIPARC:UPI0000177388

A;Note: the authors translated the codon GAT for residue 548 as Glu  
R;Yamada, Y.; Mudryj, M.; Sullivan, M.; de Crombrughe, B.  
J. Biol. Chem. 258, 2758-2761, 1983  
A;Title: Isolation and characterization of a genomic clone encoding chick alpha type II  
A;Reference number: A20855; MUID:83135706; PMID:6298201  
A;Accession: A20855  
A;Molecule type: DNA  
A;Residues: 222-233, 'R', 235-236, 'V', 238-239; 330-336, 'V', 338-351 <YAM>  
A;Cross-references: UNIPARC:UPI00000FB63A; UNIPARC:UPI0000177389  
A;Note: the authors translated the codon CGT for residue 234 as Gly, GTT for residue 277  
C;Comment: Chicken collagen alpha 1(III) chain has about 50 exons. This sequence corres  
9 (240-275), 6 (276-293), 5 (294-329), 4 (330-423), 3 (424-484), 2 (485-567), and 1 (568-615)  
C;Genetics:  
A;Introns: 68/3; 86/3; 119/3; 134/3; 152/3; 188/3; 221/3; 239/3; 275/3; 293/3; 423/3; 484/3;  
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
C;Keywords: coiled coil; extracellular matrix; glycoprotein; tandem repeat; trimer; trip  
F;1-68/Region: amino-terminal propeptide nonhelical  
F;2-62/Domain: von Willebrand factor type C repeat homology <WVC>  
F;69-347/Domain: collagen alpha helical chain (fragments) #status predicted <CAH>  
F;348-615/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
F;387-615/Domain: fibrillar collagen carboxyl-terminal homology <FCC>  
F;347,348/Disulfide bonds: interchain #status experimental  
F;516/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 24.7%; Score 58; DB 2; Length 615;  
Best Local Similarity 40.0%; Pred. No. 19;  
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 8 DTWLPKCKSLCKWHGQLRC 27  
Db 14 DWRPEPCQICVDSGSVLC 33

RESULT 13  
I50694  
collagen alpha 1(III) chain - chicken (fragment)  
C;Species: Gallus gallus (chicken)  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C;Accession: I50694  
R;Nah, H.D.; Niu, Z.; Adams, S.L.  
J. Biol. Chem. 269, 16443-16448, 1994  
A;Title: An alternative transcript of the chick type III collagen gene that does not enc  
A;Reference number: A54041; MUID:94266842; PMID:8206952  
A;Accession: I50694  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-886 <NAH>  
A;Cross-references: UNIPROT:P12105; UNIPARC:UPI000017125A; EMBL:U07973; NID:G520454; PID  
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
F;30-90/Domain: von Willebrand factor type C repeat homology <WVC>

Query Match 24.7%; Score 58; DB 2; Length 886;  
Best Local Similarity 40.0%; Pred. No. 25;  
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 8 DTWLPKCKSLCKWHGQLRC 27  
Db 42 DWRPEPCQICVDSGSVLC 61

RESULT 14  
S59856  
collagen alpha 1(III) chain precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 10-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 09-Jul-2004  
C;Accession: S59856; S62120; S16373  
R;Toman, P.D.; de Crombrughe, B.  
Gene 147, 161-168, 1994  
A;Title: The mouse type-III procollagen-encoding gene: genomic cloning and complete DNA  
A;Reference number: S59856; MUID:95011609; PMID:7926795  
A;Accession: S59856



Biochemistry 19, 1593-1599, 1980  
A;Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB5 from ty  
A;Reference number: A90438; MUID:80198282; PMID:6246925  
A;Accession: A90438  
A;Molecule type: protein  
A;Residues: 728-895, 'A', 897-964 <SEY4>  
A;Cross-references: UNIPARC:UPI00000173B85  
A;Experimental source: liver  
R;Cole, W.G.; Chiodo, A.A.; Lamande, S.R.; Janeczko, R.; Ramirez, F.; Dahl, H.H.M.; Chan  
J. Biol. Chem. 265, 17070-17077, 1990  
A;Title: A base substitution at a splice site in the COL3A1 gene causes exon skipping an  
A;Reference number: A38303; MUID:91009133; PMID:2145268  
A;Accession: A38303  
A;Molecule type: mRNA  
A;Residues: 861-1015 <COL>  
A;Cross-references: UNIPARC:UPI00000004A1; GB:J05617; GB:M55603; GB:M59227; NID:g180878;  
A;Note: a mutant sequence with 942-977 spliced out from a patient with Ehlers-Danlos syn  
R;Mankoo, B.S.; Dalglish, R.  
Nucleic Acids Res. 16, 2337, 1988  
A;Title: Human pro alpha1(III) collagen: cDNA sequence for the 3' end.  
A;Reference number: S02119; MUID:88189827; PMID:3357782  
A;Accession: S02119  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 950-1018, 'Y', 1020-1183, 'S', 1185-1466 <MAN>  
A;Cross-references: UNIPARC:UPI0000173B86; EMBL:X06700; NID:G30053; PIDN:CAA29886.1; PID  
R;Seyer, J.M.; Kang, A.H.  
Biochemistry 20, 2621-2627, 1981  
A;Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB9 from ty  
A;Reference number: A90446; MUID:81208139; PMID:7016180  
A;Accession: A90446  
A;Molecule type: protein  
A;Residues: 965-979, 'A', 981-984, 'PS', 987, 'QN', 990-1096, 'P', 1098-1152, 'AT', 1155, 'S', 1157-  
A;Cross-references: UNIPARC:UPI0000173B87  
A;Experimental source: liver  
R;Loid, H.R.; Brinker, J.M.; May, M.; Pihlajaniemi, T.; Morrow, S.; Rosenbloom, J.; Mye  
Nucleic Acids Res. 12, 9383-9394, 1984  
A;Title: Molecular cloning and carboxyl-propeptide analysis of human type III procollage  
A;Reference number: A93551; MUID:85087944; PMID:6096827  
A;Accession: A93551  
A;Molecule type: mRNA  
A;Residues: 1065-1155, 'P', 1157-1466 <LOI>  
A;Cross-references: UNIPARC:UPI0000173B88; EMBL:X01655; EMBL:X01742; NID:Q29584; PIDN:CA  
R;Miskulin, M.; Dalglish, R.; Kluge-Becker, B.; Rennard, S.I.; Tolstoshev, P.; Brant  
Biochemistry 25, 1408-1413, 1986  
A;Title: Human type III collagen gene expression is coordinately modulated with the type  
A;Reference number: 152393; MUID:86187804; PMID:3754462  
A;Accession: 152393  
A;Status: translated from GB/EMBL/DBD  
A;Molecule type: mRNA  
A;Residues: 1161-1200 <MIS>  
A;Cross-references: UNIPARC:UPI000016A6B5; GB:M13146; NID:g180415; PIDN:AAA52003.1; PID:  
R;Emanuel, B.S.; Cannizzaro, L.A.; Seyer, J.N.; Myers, J.C.  
Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985  
A;Title: Human alpha 1(III) and alpha 2(V) procollagen genes are located on the long arm  
A;Reference number: 159025; MUID:85216505; PMID:3858826  
A;Accession: 179359  
A;Status: translated from GB/EMBL/DBD  
A;Molecule type: mRNA  
A;Residues: 1165-1196 <ENA>  
A;Cross-references: UNIPARC:UPI000016A6B6; GB:M11134; NID:g180417; PIDN:AAA52004.1; PID:  
R;Chu, M.L.; Weil, D.; de Wet, W.; Bernard, M.; Sippola, M.; Ramirez, F.  
J. Biol. Chem. 260, 4357-4363, 1985  
A;Title: Isolation of cDNA and genomic clones encoding human pro-alpha1(III) collagen. F  
A;Reference number: A92516; MUID:85157600; PMID:2579949  
A;Accession: A92516  
A;Molecule type: DNA  
A;Residues: 1176-1240, 'V', 1242-1356, 'P', 1358-1466 <CHU>  
A;Cross-references: UNIPARC:UPI0000173B89; GB:M10793; GB:M10794; GB:M10795; C  
A;Experimental source: liver  
A;Note: the authors translated the codon TTC for residue 1057 as Tyr; the codons given f  
ation  
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit (

3-hydroxylated. About 15% of the lysines are 5-hydroxylated and some are subsequently O  
C;Genetics:  
A;Gene: GDB:COL3A1  
A;Cross-references: GDB:118729; OMIM:120180  
A;Map position: 2q31-2q31  
A;Introns: 27/1; 94/3; 111/3; 149/3; 176/3; 554/3; 587/3; 1175/3; 1275/3; 1337/3; 1418/3;  
A;Note: the list of introns is incomplete; defects in this gene can result in Ehlers-Dan  
C;Complex: type III collagen is a homotrimer of monomers initially linked by disulfide b  
er of their length, is formed with desmosine cross-links made from lysine and allysine r  
C;Function:  
A;Description: structural component of extracellular fibrous polymer that maintains inte  
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
C;Keywords: coiled coil; Ehlers-Danlos syndrome; extracellular matrix; glycoprotein; hy  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-153/Domain: amino-terminal propeptide #status predicted <PRC>  
F;31-91/Domain: von Willebrand factor type C repeat homology <VWC>  
F;154-1221/Product: collagen alpha 1(III) chain #status predicted <MAT>  
F;154-167/Region: amino-terminal nonhelical telopeptide  
F;168-1196/Region: helical  
F;1091-1093/Region: cell attachment (R-G-D) motif  
F;1197-1221/Region: carboxyl-terminal nonhelical telopeptide  
F;1222-1466/Domain: carboxyl-terminal propeptide #status predicted <CPR>  
F;1238-1466/Domain: fibrillar collagen carboxyl-terminal homology <FCC>  
F;24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F;153-154/Cleavage site: Pro-Gln (procollagen N-endopeptidase) #status predicted  
F;154/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F;161,1212/Modified site: allysine (Lys) #status predicted

Query Match 24.78; Score 58; DB 1; Length 1466;  
Best Local Similarity 40.0%; Pred. No. 39;  
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 8 DTWLPKCKSLCKWHGQLRC 27

Db 43 DVWKPEPCQICVDSGSVLC 62

Search completed: February 3, 2006, 18:31:17  
Job time : 7.08444 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:15:25 ; Search time 34.78 Seconds  
(without alignments)  
467.424 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_114\_150

Perfect score: 235

Sequence: 1 NCSPVPHDTWLPKCSLCCKWHGQLRCPQAFPLPGCD 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*
- 9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	235	100.0	129	5	ABB77111 Human Cri
2	235	100.0	129	5	ABB77124 Human Cri
3	235	100.0	129	5	ABB77104 Human Cri
4	235	100.0	129	5	ABB77117 Human Cri
5	235	100.0	139	5	ABB77103 Human Cri
6	235	100.0	139	5	ABB77123 Human Cri
7	235	100.0	139	5	ABB77110 Human Cri
8	235	100.0	139	5	ABB77116 Human Cri
9	235	100.0	173	8	ADS88698 Amino aci
10	235	100.0	174	2	AAR13326 Recombina
11	235	100.0	174	2	AAW32107
12	235	100.0	174	8	ADO05065 Human cri
13	235	100.0	183	8	ADS88699 Amino aci
14	235	100.0	184	8	ADO28629 Human Cri
15	235	100.0	188	2	AAR22548 Human Cri
16	235	100.0	188	2	AAR22547 Human Cri
17	235	100.0	188	2	AAW25667 Protein e
18	235	100.0	188	2	AAW32108 Human Cri
19	235	100.0	188	2	AAW19980 Human Cri
20	235	100.0	188	2	AAW29735 Homo sapi
21	235	100.0	188	2	AAW87631 Human Cri
22	235	100.0	188	2	AAW87630 Human Cri
23	235	100.0	188	5	ABB77102 Human Cri
24	235	100.0	188	5	ABB77101 Human Cri

25	235	100.0	188	5	ABB77122	Abb77122 Human Cri
26	235	100.0	188	5	ABB77109	Abb77109 Human Cri
27	235	100.0	188	5	AAO14637	AAO14637 Human Cri
28	235	100.0	188	5	AAO14638	AAO14638 Human Cri
29	235	100.0	188	5	AAO14636	AAO14636 Human Cri
30	235	100.0	188	5	AAO14727	AAO14727 Human var
31	235	100.0	188	6	ABP97176	ABP97176 Tumour-as
32	235	100.0	188	6	ABP58132	ABP58132 Human Cri
33	235	100.0	188	6	ABP58131	ABP58131 Human Cri
34	235	100.0	188	7	ADC78799	ADC78799 Human PRO
35	235	100.0	188	7	ADC78843	ADC78843 Human PRO
36	235	100.0	188	7	ADD26059	ADD26059 CR-2 crip
37	235	100.0	188	7	ADD26058	ADD26058 CR-1 crip
38	235	100.0	188	8	ADI82176	ADI82176 Human ter
39	235	100.0	188	8	ADR70486	ADR70486 Human ter
40	235	100.0	188	8	ADS88697	ADS88697 Amino aci
41	235	100.0	188	8	ADT79198	ADT79198 Human Cri
42	235	100.0	188	8	ADT79200	ADT79200 Human Cri
43	235	100.0	188	8	ADT79197	ADT79197 Human Cri
44	235	100.0	188	8	ADT79290	ADT79290 Human Cri
45	235	100.0	188	9	ADY85961	ADY85961 Human Cri

#### ALIGNMENTS

RESULT 1  
ABB77111  
ID ABB77111 standard; protein; 129 AA.  
XX  
AC ABB77111;  
XX  
DT 08-OCT-2002 (first entry)  
XX  
DE Human Cripto-1 C-terminal truncated form mutant CR38(cc)T88A.  
XX  
KW Human; Cripto-1; CR-1; mutant; tumour; cell proliferation.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 48  
FT /note= "Wild type Thr substituted by Ala"  
XX  
PN WO200222808-A2.  
XX  
PD 21-MAR-2002.  
XX

18-SEP-2001; 2001WO-US029066.  
18-SEP-2000; 2000US-0233148P.  
(BIOJ ) BIOGEN INC.  
Williams KP, Foley S, Schiffer S, Doman B, Sanicola-Nadel M; WPI; 2002-339868/37.  
New mutant form of CRYPTO (teratocarcinoma-derived growth factor), useful for treating cell proliferation, especially cancer, comprises amino acid change that prevents fucosylation at Thr88.  
Disclosure; Page 36; 41pp; English.

The sequence represents a C-terminal truncated form of human Cripto-1, where the threonine residue at position 48 has been substituted with an alanine residue. The invention relates to a novel mutant CRYPTO polypeptide, or its functional fragment, having at least one amino acid alteration at positions 86, 87 or 88. The mutant polypeptide, or its chimera, is used to inhibit growth of tumour cells, in vivo or in vitro, particularly for treating breast, ovarian, renal, colorectal, uterine, prostatic, lung, bladder or central nervous system cancers, melanoma and

CC leukaemia, also generally for treating undesired cell proliferation

XX

SQ Sequence 129 AA;

Query Match 100.0%; Score 235; DB 5; Length 129;  
Best Local Similarity 100.0%; Pred. No. 5.4e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37  
|||||  
74 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 110

Db

RESULT 2

ABB77124

ID ABB77124 standard; protein; 129 AA.

XX

AC ABB77124;

XX

08-OCT-2002 (first entry)

DE Human Cripto-3 C-terminal truncated form mutant CR-338(cc)T88A.

XX

Human Cripto-1; CR-1; mutant; tumour; Cripto-3; CR-3;  
cell proliferation.

XX

Homo sapiens.

OS Synthetic.

XX

Key Location/Qualifiers

FT Misc-difference 48 /note= "wild type Thr substituted by Ala"

XX

WO200222808-A2.

XX

21-MAR-2002.

XX

18-SEP-2001; 2001WO-US029066.

XX

18-SEP-2000; 2000US-0233148P.

XX

(BIOJ ) BIOGEN INC.

XX

Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
WPI; 2002-339868/37.

XX

New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful  
for treating cell proliferation, especially cancer, comprises amino acid  
change that prevents fucosylation at Thr88.

XX

Disclosure; Page 39; 41pp; English.

XX

The sequence represents a C-terminal truncated form of human Cripto-3,  
where the threonine residue at position 48 has been substituted with an  
alanine residue. The invention relates to a novel mutant CRIPTO  
polypeptide, or its functional fragment, having at least one amino acid  
alteration at positions 86, 87 or 88. The mutant polypeptide, or its  
chimer, is used to inhibit growth of tumour cells, in vivo or in vitro,  
particularly for treating breast, ovarian, renal, colorectal, uterine,  
prostatic, lung, bladder or central nervous system cancers, melanoma and  
leukaemia, also generally for treating undesired cell proliferation

XX

SQ Sequence 129 AA;

Query Match 100.0%; Score 235; DB 5; Length 129;  
Best Local Similarity 100.0%; Pred. No. 5.4e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37  
|||||  
74 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 110

Db

RESULT 3

ABB77104

ID ABB77104 standard; protein; 129 AA.

XX

AC ABB77104;

XX

08-OCT-2002 (first entry)

DE Human Cripto-1 C-terminal truncated form (CR38(cc)).

XX

Human; Cripto-1; CR-1; mutant; tumour; cell proliferation.

XX

Homo sapiens.

OS WO200222808-A2.

XX

21-MAR-2002.

XX

18-SEP-2001; 2001WO-US029066.

XX

18-SEP-2000; 2000US-0233148P.

XX

(BIOJ ) BIOGEN INC.

XX

Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
WPI; 2002-339868/37.

XX

New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful  
for treating cell proliferation, especially cancer, comprises amino acid  
change that prevents fucosylation at Thr88.

XX

Claim 2; Page 35; 41pp; English.

XX

The sequence represents a C-terminal truncated form of human Cripto-1.  
The invention relates to a novel mutant CRIPTO polypeptide, or its  
functional fragment, having at least one amino acid alteration at  
positions 86, 87 or 88. The mutant polypeptide, or its chimera, is used  
to inhibit growth of tumour cells, in vivo or in vitro, particularly for  
treating breast, ovarian, renal, colorectal, uterine, prostatic, lung,  
bladder or central nervous system cancers, melanoma and leukaemia, also  
generally for treating undesired cell proliferation

XX

SQ Sequence 129 AA;

Query Match 100.0%; Score 235; DB 5; Length 129;  
Best Local Similarity 100.0%; Pred. No. 5.4e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 37  
|||||  
74 NCGSVPHDTWLPKCKSLCKWHGQLRCFPQAFPLPGCD 110

Db

RESULT 4

ABB77117

ID ABB77117 standard; protein; 129 AA.

XX

AC ABB77117;

XX

08-OCT-2002 (first entry)

DE Human Cripto-3 C-terminal truncated form (CR-338(cc)).

XX

Human; Cripto-1; CR-1; mutant; tumour; Cripto-3; CR-3;  
cell proliferation.

XX

Homo sapiens.

OS WO200222808-A2.

XX

21-MAR-2002.

XX





```
QY      1 NCGSVPHDTWLPKCKSLCKCWHGQLRCFPQAFPLPGCD 37
DB      84 NCGSVPHDTWLPKCKSLCKCWHGQLRCFPQAFPLPGCD 120

RESULT 7
ABB77110
ID  ABB77110 standard; protein; 139 AA.
XX  AC
XX  AC
XX  DT 08-OCT-2002 (first entry)
XX  DE Human Cripto-1 C-terminal truncated form mutant CR38(cc)T88A.
XX  KW Human; Cripto-1; CR-1; mutant; tumour; cell proliferation.
XX  OS Homo sapiens.
XX  OS Synthetic.
XX  FH Key Location/Qualifiers
XX  FT Misc-difference 58
XX  FT /note= "Wild type Thr substituted by Ala"
XX  XX
XX  PN WO200222808-A2.
XX  PD 21-MAR-2002.
XX  PF 18-SEP-2001; 2001WO-US029066.
XX  PR 18-SEP-2000; 2000US-0233148P.
XX  PA (BIOJ ) BIOGEN INC.
XX  PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;
XX  PI WPI; 2002-339868/37.
XX  DR
XX  PT New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful
XX  PT for treating cell proliferation, especially cancer, comprises amino acid
XX  PT change that prevents fucosylation at Thr88.
XX  PS Example 3; Page 36; 41pp; English.
XX  CC The sequence represents a mutant C-terminal truncated form of human
XX  CC Cripto-1, where the threonine residue at position 58 has been replaced
XX  CC with an alanine residue. The invention relates to a novel mutant CRIPTO
XX  CC polypeptide, or its functional fragment, having at least one amino acid
XX  CC alteration at positions 86, 87 or 88. The mutant polypeptide, or its
XX  CC chimera, is used to inhibit growth of tumour cells, in vivo or in vitro,
XX  CC particularly for treating breast, ovarian, renal, colorectal, uterine,
XX  CC prostatic, lung, bladder or central nervous system cancers, melanoma and
XX  CC leukaemia, also generally for treating undesired cell proliferation
XX  SQ Sequence 139 AA;

Query Match      100.0%; Score 235; DB 5; Length 139;
Best Local Similarity 100.0%; Pred. NO. 5.8e-20;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Example 3; Page 36; 41pp; English.

The sequence represents a mutant C-terminal truncated form of human
Cripto-1, where the threonine residue at position 58 has been replaced
with an alanine residue. The invention relates to a novel mutant CRIPTO
polypeptide, or its functional fragment, having at least one amino acid
alteration at positions 86, 87 or 88. The mutant polypeptide, or its
chimera, is used to inhibit growth of tumour cells, in vivo or in vitro,
particularly for treating breast, ovarian, renal, colorectal, uterine,
prostatic, lung, bladder or central nervous system cancers, melanoma and
leukaemia, also generally for treating undesired cell proliferation

Query Match      100.0%; Score 235; DB 5; Length 139;
Best Local Similarity 100.0%; Pred. NO. 5.8e-20;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NCGSVPHDTWLPKCKSLCKCWHGQLRCFPQAFPLPGCD 37
DB      84 NCGSVPHDTWLPKCKSLCKCWHGQLRCFPQAFPLPGCD 120

RESULT 8
ABB77116
ID  ABB77116 standard; protein; 139 AA.
XX  AC
XX  AC
XX  DT 08-OCT-2002 (first entry)
```

```
XX      Human Cripto-3 C-terminal truncated form (CR-3(cc)).
DE
XX      Human; Cripto-1; CR-1; mutant; tumour; Cripto-3; CR-3;
KW      cell proliferation.
XX
XX      Homo sapiens.
OS
XX      WO200222808-A2.
PN
XX      21-MAR-2002.
XX
XX      18-SEP-2001; 2001WO-US029066.
XX
XX      18-SEP-2000; 2000US-0233148P.
PR
XX      (BIOJ ) BIOGEN INC.
XX
XX      Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;
PI
XX      WPI; 2002-339868/37.
DR
XX      New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful
PT      for treating cell proliferation, especially cancer, comprises amino acid
PT      change that prevents fucosylation at Thr88.
XX
XX      Claim 2; Page 37; 41pp; English.
XX
XX      The sequence represents a C-terminal truncated form of human Cripto-3.
CC      The invention relates to a novel mutant CRIPTO polypeptide, or its
CC      functional fragment, having at least one amino acid alteration at
CC      positions 86, 87 or 88. The mutant polypeptide, or its chimera, is used
CC      to inhibit growth of tumour cells, in vivo or in vitro, particularly for
CC      treating breast, ovarian, renal, colorectal, uterine, prostatic, lung,
CC      bladder or central nervous system cancers, melanoma and leukaemia, also
CC      generally for treating undesired cell proliferation
XX
XX      SQ Sequence 139 AA;

Query Match      100.0%; Score 235; DB 5; Length 139;
Best Local Similarity 100.0%; Pred. NO. 5.8e-20;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NCGSVPHDTWLPKCKSLCKCWHGQLRCFPQAFPLPGCD 37
DB      84 NCGSVPHDTWLPKCKSLCKCWHGQLRCFPQAFPLPGCD 120

RESULT 9
ADS88698
ID  ADS88698 standard; protein; 173 AA.
XX
XX      ADS88698;
AC
XX      16-DEC-2004 (first entry)
DT
XX
XX      Amino acid sequence of a human secreted Cripto protein.
DE
XX
XX      stem cell differentiation; cardiomyocyte; stem cell;
KW      epidermal growth factor; EGF; EGF-CFC family; Cripto; neuronal cell;
KW      Alq4; Nodal; heart disease; neuropathology; myocardial infarction;
KW      Parkinson's disease; Alzheimer's disease; retinal degeneration;
KW      cell therapy.
XX
XX      Homo sapiens.
OS
XX      WO2004083375-A2.
PN
XX      30-SEP-2004.
XX
XX      19-MAR-2004; 2004WO-IT000133.
PF
XX      21-MAR-2003; 2003IT-RM000125.
PR
```



CC colorectal and breast carcinoma. Recombinant CRIPTO can be expressed in  
CC prokaryotic or eukaryotic host cells, and used for the identification and  
CC characterisation of specific cell surface receptors, or to raise specific  
CC antibodies to assay CRIPTO protein expression in immunoassays. (Updated  
CC on 25-MAR-2003 to correct PF field.)  
XX  
XX  
SQ Sequence 174 AA;

Query Match 100.0%; Score 235; DB 2; Length 174;  
Best Local Similarity 100.0%; Pred. No. 7.1e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKKGSLCKWHGQLRCFPQAFPLPGCD 37  
|||||  
Db 100 NCGSVPHDTWLPKKGSLCKWHGQLRCFPQAFPLPGCD 136  
|||||

## RESULT 12

ADO05065  
ID ADO05065 standard; protein; 174 AA.

XX ADO05065;

XX 29-JUL-2004 (first entry)

XX Human cripto growth factor protein.

XX Criptin growth factor; CGF; wound healing; tissue regeneration;  
KW implant fixation; angiogenesis; neoplasia; tumour; gene therapy; human;  
KW cripto growth factor.  
XX

OS Homo sapiens.

XX US2004086967-A1.

XX 06-MAY-2004.

XX 22-SEP-2003; 2003US-00665602.

XX 06-JUN-1995; 95US-00471371.

XX 09-SEP-1999; 99US-00393023.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Meissner PS, Coleman TA;

XX WPI; 2004-356201/33.

XX New human polynucleotides encoding human criptin growth factor  
PT polypeptides, useful for wound healing or tissue regeneration,  
PT stimulating implant fixation and angiogenesis, and for treating and/or  
PT preventing tumor.

XX Disclosure; SEQ ID NO 7; 19pp; English.

XX The invention provides criptin growth factor (CGF) polypeptides and their  
CC encoding polynucleotides. The invention is useful for wound healing and  
CC tissue regeneration, stimulating implant fixation, angiogenesis and for  
CC treating and preventing neoplasia such as tumour. The invention is also  
CC useful in gene therapy. The present sequence is human cripto growth  
CC factor protein. This sequence is used in the invention.  
XX

XX Sequence 174 AA;

Query Match 100.0%; Score 235; DB 8; Length 174;  
Best Local Similarity 100.0%; Pred. No. 7.1e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKKGSLCKWHGQLRCFPQAFPLPGCD 37  
|||||  
Db 100 NCGSVPHDTWLPKKGSLCKWHGQLRCFPQAFPLPGCD 136  
|||||

## RESULT 13

ADS88699  
ID ADS88699 standard; protein; 183 AA.

XX ADS88699;

XX 16-DEC-2004 (first entry)

XX Amino acid sequence of a human secreted His-tagged Cripto protein.

XX stem cell differentiation; cardiomyocyte; stem cell;

XX epidermal growth factor; EGF; EGF-CFC family; Cripto; neuronal cell;

XX Alq4; Nodal; heart disease; neuropathology; myocardial infarction;

XX Parkinson's disease; Alzheimer's disease; retinal degeneration;

XX cell therapy.

XX Homo sapiens.

XX Synthetic.

XX Key Location/Qualifiers

XX FT 174..183

XX /note= "His tag"

XX WO2004083375-A2.

XX 30-SEP-2004.

XX 19-MAR-2004; 2004WO-IT000133.

XX 21-MAR-2003; 2003IT-RM000125.

XX 29-JUL-2003; 2003IT-RM000370.

XX (MNC/) MINCHIOTTI G.

XX (PERS/) PERSICO M.

XX (PARI/) PARISI S.

XX Minchiotti G, Persico M, Parisi S;

XX WPI; 2004-691035/67.

XX Inducing stem cell differentiation in cardiomyocytes or neuronal cells by  
PT exposure to a protein of EGF-CFC or Cripto protein inhibitors,  
PT respectively, useful in treating heart diseases and neuropathologies.

XX Disclosure; Page 13; 35pp; English.

XX The specification describes a method for inducing stem cell  
CC differentiation into cardiomyocytes. The method comprises exposing the  
CC cells to a protein of the epidermal growth factor (EGF)-CFC family or its  
CC derivatives and having at least the EGF and CFC domains (e.g. Cripto  
protein). A second method is described for stem cell differentiation into  
CC neuronal cells, where the cells are exposed to an inhibitor of the Cripto  
protein. The Cripto protein inhibitor is an anti-Cripto antibody or its  
CC functional fragments, or is a peptide specifically selected from a random  
CC combinatorial peptide library, or is an antagonist of the Alq4 (receptor)-  
CC Cripto(co-receptor)-Nodal(ligand) pathway. The differentiated stem cells  
CC of the invention are useful in the treatment of heart diseases and  
CC neuropathologies, e.g. myocardial infarction, Parkinson's disease,  
CC Alzheimer's disease and retinal degeneration. They are especially useful  
CC for cell therapy. The Cripto protein or its inhibitors is useful in the  
CC preparation of a composition able to direct stem cell differentiation  
CC toward the neuronal lineage. The present sequence represents a Cripto  
CC protein, which maybe used in the method of the invention.

XX Sequence 183 AA;

Query Match 100.0%; Score 235; DB 8; Length 183;  
Best Local Similarity 100.0%; Pred. No. 7.5e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NCGSVPHDTWLPKKGSLCKWHGQLRCFPQAFPLPGCD 37  
|||||  
Db 114 NCGSVPHDTWLPKKGSLCKWHGQLRCFPQAFPLPGCD 150  
|||||

RESULT 14  
ADO28629  
ID ADO28629 standard; protein; 184 AA.  
XX AC ADO28629;  
XX DT 12-AUG-2004 (first entry)  
XX DE Human CRIPTO CR-1 protein SEQ ID NO:58.  
XX KW high-grade dysplasia; HGD; oesophageal adenocarcinoma;  
XX KW neo-plastic transformation; cancer; cytostatic; gene therapy; human;  
XX KW CRIPTO CR-1; chromosome 3.  
XX OS Homo sapiens.  
XX PN WO2004044178-A2.  
XX PD 27-MAY-2004.  
XX PF 13-NOV-2003; 2003WO-US036260.  
XX PR 13-NOV-2002; 2002US-0425813P.  
XX PA (GETH ) GENENTECH INC.  
XX PI Smith V;  
XX DR WPI; 2004-420319/39.  
XX DR N-PSDB; ADO28628.  
XX PT Detecting of high-grade dysplasia in cells of a mammalian tissue sample  
XX PT comprises establishing the level of expression in the test tissue sample  
XX PT of the genes.  
XX PS Example 4; SEQ ID NO 58; 256pp; English.  
XX CC The present invention describes a method for detecting high-grade  
XX CC dysplasia (HGD) in cells of a mammalian tissue sample. Also described:  
XX CC (1) identifying an oesophageal tissue susceptible to oesophageal  
XX CC adenocarcinoma; (2) determining the predisposition of a mammalian tissue  
XX CC to a neo-plastic transformation by detecting HGD in cells of the tissue;  
XX CC and (3) detecting cancer in a patient. The method can be used in  
XX CC detecting HGD and cancer in cells of a mammalian tissue sample. The  
XX CC methods and compositions of the present invention can be used in treating  
XX CC and preventing HGD and cancer, and in gene therapy. The present sequence  
XX CC represents human CRIPTO CR-1, which is used in the exemplification of the  
XX CC present invention. The human CRIPTO CR-1 gene is located on chromosome 3.  
XX SQ Sequence 184 AA;  
Query Match 100.0%; Score 235; DB 8; Length 184;  
Best Local Similarity 100.0%; Pred. No. 7.5e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NCGSVPHDWTLPKCKSLCKWHGQLRCFPOAFLPGCD 37  
DB 111 NCGSVPHDWTLPKCKSLCKWHGQLRCFPOAFLPGCD 147  
RESULT 15  
AAR22548  
ID AAR22548 standard; protein; 188 AA.  
XX AC AAR22548;  
XX DT 25-MAR-2003 (revised)  
XX DT 17-DEC-2001 (revised)  
XX DT 28-AUG-1992 (first entry)  
XX DE Human CRIPTO CR-1 protein.

XX cell proliferation; tumour; CR-3; transforming growth factor;  
KW epidermal growth factor.  
XX OS Homo sapiens.  
XX PN USN7749001-N.  
XX PD 25-FEB-1992.  
XX PF 23-AUG-1991; 91US-00749001.  
XX PR 23-AUG-1991; 91US-00749001.  
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
XX PI Salomom D, Persico M;  
XX DR WPI; 1992-123675/15.  
XX DR N-PSDB; AAQ22495.  
XX PT New crypto gene CR-1 and crypto-related gene CR-3 genomic DNA - CR-3  
XX PT protein and anti-CR-3 antibodies, useful in immunoassay to detect CR-3 as  
XX PT tumour specific marker.  
XX PS Disclosure; Page 22; 44pp; English.  
XX CC The CR-1 genomic clone which codes for this protein can be used in  
XX CC transgenic animals to examine the effects of overexpression of this gene  
XX CC on development and tumorigenicity and to study regulation of CR-1 gene.  
XX CC See also AAQ22494. (Note: Revised entry submitted to correct the patent  
XX CC number format of US Government-Owned NTIS applications to prevent clashes  
XX CC with ongoing US granted patent numbers. For further information please  
XX CC visit the Derwent web site at [www.derwent.com/dwpi/updates/ntis\\_us.html](http://www.derwent.com/dwpi/updates/ntis_us.html).)  
XX CC (Updated on 25-MAR-2003 to correct PA field.)  
XX SQ Sequence 188 AA;  
Query Match 100.0%; Score 235; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 7.7e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NCGSVPHDWTLPKCKSLCKWHGQLRCFPOAFLPGCD 37  
DB 114 NCGSVPHDWTLPKCKSLCKWHGQLRCFPOAFLPGCD 150  
Search completed: February 3, 2006, 18:22:37  
Job time : 34.78 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:58:17 ; Search time 10 Seconds

(without alignments)  
19.921 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_46\_62

Perfect score: 96

Sequence: 1 FRDSIWQEEPAIRPR 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

Published Applications\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pap:\*  
2: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pap:\*  
3: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pap:\*  
4: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pap:\*  
5: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pap:\*  
6: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pap:\*  
7: /cgn2\_6/ptodata/2/pubppaa/US11\_NEW\_PUB.pap:\*  
8: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	17	6	US-10-945-853-3
2	96	100.0	188	6	US-10-945-853-1
3	96	100.0	188	6	US-10-945-853-2
4	44	45.8	115	7	US-11-000-463-252
5	40	41.7	107	7	US-11-064-174-29
6	39.5	41.1	228	6	US-10-467-657-156
7	39.5	41.1	255	6	US-10-467-657-6148
8	39.5	41.1	255	6	US-10-467-657-7562
9	39	40.6	305	6	US-10-520-820-1
10	39	40.6	441	6	US-10-055-877-4
11	39	40.6	441	6	US-10-055-877-135
12	39	40.6	479	6	US-10-055-877-134
13	39	40.6	485	6	US-10-055-877-133
14	39	40.6	637	6	US-10-055-877-131
15	39	40.6	774	7	US-11-077-886-34
16	39	40.6	988	7	US-11-171-701-6
17	38	39.6	17	6	US-10-503-575-83
18	38	39.6	92	6	US-10-467-657-7960
19	38	39.6	335	7	US-11-092-353-2
20	38	39.6	478	7	US-11-092-353-4
21	38	39.6	2050	6	US-10-453-372-192
22	38	39.6	3588	6	US-10-453-372-194
23	38	39.6	3570	6	US-10-453-372-178
24	38	39.6	3570	6	US-10-453-372-196
25	38	39.6	3570	6	US-10-453-372-198

Sequence 200, App  
Sequence 202, App  
Sequence 204, App  
Sequence 206, App  
Sequence 56, Appl  
Sequence 153, App  
Sequence 26, Appl  
Sequence 28, Appl  
Sequence 301, App  
Sequence 386, App  
Sequence 286, App  
Sequence 31, App  
Sequence 83, Appl  
Sequence 105, App  
Sequence 93, Appl  
Sequence 105, App  
Sequence 69, Appl  
Sequence 67, Appl  
Sequence 869, App  
Sequence 1505, App  
Sequence 195, App  
Sequence 9554, App  
Sequence 4, Appl  
Sequence 58, Appl  
Sequence 20, Appl  
Sequence 28, Appl  
Sequence 20, Appl  
Sequence 28, Appl  
Sequence 1204, App  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 262, App  
Sequence 51, Appl  
Sequence 31, Appl  
Sequence 3488, App  
Sequence 8180, App  
Sequence 8176, App  
Sequence 142, App  
Sequence 142, App  
Sequence 298, App  
Sequence 12, Appl  
Sequence 138, App  
Sequence 154, App  
Sequence 253, App  
Sequence 8730, App  
Sequence 296, App  
Sequence 388, App  
Sequence 179, App  
Sequence 598, App  
Sequence 696, App  
Sequence 1328, App  
Sequence 45, Appl  
Sequence 5524, App  
Sequence 252, App  
Sequence 59, Appl  
Sequence 71, Appl  
Sequence 116, App  
Sequence 123, App  
Sequence 65, Appl  
Sequence 116, App  
Sequence 116, App  
Sequence 123, App  
Sequence 40, Appl  
Sequence 40, Appl  
Sequence 6, Appl

3570 39.6 38 26  
3570 39.6 38 27  
3570 39.6 38 28  
3570 39.6 38 29  
3570 39.6 38 30  
3570 39.6 38 31  
3570 39.6 38 32  
3570 39.6 38 33  
3570 39.6 38 34  
3570 39.6 38 35  
3570 39.6 38 36  
3570 39.6 38 37  
3570 39.6 38 38  
3570 39.6 38 39  
3570 39.6 38 40  
3570 39.6 38 41  
3570 39.6 38 42  
3570 39.6 38 43  
3570 39.6 38 44  
3570 39.6 38 45  
3570 39.6 38 46  
3570 39.6 38 47  
3570 39.6 38 48  
3570 39.6 38 49  
3570 39.6 38 50  
3570 39.6 38 51  
3570 39.6 38 52  
3570 39.6 38 53  
3570 39.6 38 54  
3570 39.6 38 55  
3570 39.6 38 56  
3570 39.6 38 57  
3570 39.6 38 58  
3570 39.6 38 59  
3570 39.6 38 60  
3570 39.6 38 61  
3570 39.6 38 62  
3570 39.6 38 63  
3570 39.6 38 64  
3570 39.6 38 65  
3570 39.6 38 66  
3570 39.6 38 67  
3570 39.6 38 68  
3570 39.6 38 69  
3570 39.6 38 70  
3570 39.6 38 71  
3570 39.6 38 72  
3570 39.6 38 73  
3570 39.6 38 74  
3570 39.6 38 75  
3570 39.6 38 76  
3570 39.6 38 77  
3570 39.6 38 78  
3570 39.6 38 79  
3570 39.6 38 80  
3570 39.6 38 81  
3570 39.6 38 82  
3570 39.6 38 83  
3570 39.6 38 84  
3570 39.6 38 85  
3570 39.6 38 86  
3570 39.6 38 87  
3570 39.6 38 88  
3570 39.6 38 89  
3570 39.6 38 90  
3570 39.6 38 91  
3570 39.6 38 92  
3570 39.6 38 93  
3570 39.6 38 94  
3570 39.6 38 95  
3570 39.6 38 96  
3570 39.6 38 97  
3570 39.6 38 98

99	34	35.4	107	7	US-11-144-222-6	Sequence 6, Appli	172	33	34.4	637	6	US-10-821-234-961	Sequence 961, App
100	34	35.4	107	7	US-11-102-424-16	Sequence 16, Appl	173	33	34.4	663	6	US-10-467-945A-1	Sequence 1, Appli
101	34	35.4	107	7	US-11-102-424-17	Sequence 17, Appl	174	33	34.4	829	6	US-10-821-234-1179	Sequence 1179, Ap
102	34	35.4	107	7	US-11-182-343-6	Sequence 6, Appli	175	33	34.4	871	7	US-11-109-157A-10	Sequence 10, Appl
103	34	35.4	108	6	US-10-925-366A-194	Sequence 194, App	176	33	34.4	882	7	US-11-151-601-20	Sequence 20, Appl
104	34	35.4	108	6	US-10-771-257-25	Sequence 25, Appl	177	33	34.4	901	7	US-11-082-389-430	Sequence 430, App
105	34	35.4	108	7	US-11-127-677-25	Sequence 25, Appl	178	33	34.4	905	7	US-11-124-368A-307	Sequence 307, App
106	34	35.4	146	7	US-11-084-858-10	Sequence 10, Appl	179	33	34.4	928	6	US-10-841-129-4	Sequence 4, Appli
107	34	35.4	148	6	US-10-530-253-22	Sequence 22, Appl	180	33	34.4	928	6	US-10-841-129-6	Sequence 6, Appli
108	34	35.4	236	7	US-11-144-248-47	Sequence 47, Appl	181	33	34.4	968	7	US-11-000-463-281	Sequence 281, App
109	34	35.4	236	7	US-11-144-248-48	Sequence 48, Appl	182	33	34.4	1103	7	US-11-109-157A-9	Sequence 9, Appli
110	34	35.4	236	7	US-11-144-248-52	Sequence 52, Appl	183	33	34.4	1142	7	US-11-044-051-73	Sequence 73, Appli
111	34	35.4	236	7	US-11-144-222-47	Sequence 47, Appl	184	33	34.4	1167	6	US-10-942-072-6	Sequence 6, Appli
112	34	35.4	236	7	US-11-144-222-48	Sequence 48, Appl	185	33	34.4	1167	6	US-10-942-072-13	Sequence 13, Appl
113	34	35.4	236	7	US-11-144-222-52	Sequence 52, Appl	186	33	34.4	1168	6	US-10-942-072-11	Sequence 11, Appl
114	34	35.4	236	7	US-11-182-343-47	Sequence 47, Appl	187	33	34.4	1271	6	US-10-770-726-46	Sequence 46, Appl
115	34	35.4	236	7	US-11-182-343-48	Sequence 48, Appl	188	33	34.4	1823	6	US-10-995-561-988	Sequence 988, App
116	34	35.4	236	7	US-11-182-343-52	Sequence 52, Appl	189	33	34.4	2102	6	US-10-995-561-990	Sequence 990, App
117	34	35.4	242	7	US-11-054-515-1844	Sequence 1844, Ap	190	33	34.4	2108	6	US-10-995-561-989	Sequence 989, App
118	34	35.4	251	7	US-11-054-515-1217	Sequence 1217, Ap	191	33	34.4	2157	6	US-10-995-561-991	Sequence 991, App
119	34	35.4	257	6	US-10-642-272A-5	Sequence 5, Appli	192	33	34.4	2556	7	US-11-050-346-67	Sequence 67, Appl
120	34	35.4	257	6	US-10-642-272A-6	Sequence 6, Appli	193	33	34.4	3597	7	US-11-019-711-6	Sequence 6, Appli
121	34	35.4	333	6	US-10-878-556A-174	Sequence 174, App	194	32.5	33.9	74	7	US-11-116-746-15	Sequence 15, Appl
122	34	35.4	427	6	US-10-995-561-816	Sequence 816, App	195	32.5	33.9	108	6	US-10-925-366A-164	Sequence 164, App
123	34	35.4	427	6	US-10-995-561-818	Sequence 818, App	196	32.5	33.9	175	7	US-11-113-581-90	Sequence 90, Appl
124	34	35.4	565	6	US-10-793-626-2608	Sequence 2608, Ap	197	32.5	33.9	177	7	US-11-176-830-651	Sequence 651, App
125	34	35.4	645	6	US-10-793-626-1770	Sequence 1770, Ap	198	32.5	33.9	267	7	US-11-074-176-52	Sequence 52, Appl
126	34	35.4	648	6	US-10-793-626-568	Sequence 568, App	199	32.5	33.9	418	6	US-10-512-325-3	Sequence 3, Appli
127	34	35.4	802	6	US-10-510-386-2	Sequence 2, Appli	200	32.5	33.9	418	7	US-11-196-919-2	Sequence 2, Appli
128	34	35.4	926	6	US-10-841-129-2	Sequence 2, Appli	201	32.5	33.9	601	6	US-10-821-234-958	Sequence 958, App
129	34	35.4	1767	6	US-10-995-561-911	Sequence 911, App	202	32.5	33.9	638	7	US-11-150-845-24	Sequence 24, Appl
130	34	35.4	1767	6	US-10-995-561-914	Sequence 914, App	203	32.5	33.9	795	6	US-10-770-726-49	Sequence 49, Appl
131	34	35.4	1806	6	US-10-995-561-912	Sequence 912, App	204	32.5	33.9	833	7	US-11-076-187-5	Sequence 5, Appli
132	34	35.4	1806	6	US-10-995-561-915	Sequence 915, App	205	32	33.3	12	6	US-10-989-226-6	Sequence 6, Appli
133	34	35.4	1818	6	US-10-995-561-910	Sequence 910, App	206	32	33.3	12	7	US-11-066-967-33	Sequence 33, Appl
134	34	35.4	1818	6	US-10-995-561-913	Sequence 913, App	207	32	33.3	20	6	US-11-188-187A-107	Sequence 107, App
135	33.5	34.9	177	7	US-11-176-830-652	Sequence 652, App	208	32	33.3	21	6	US-10-723-207-69	Sequence 69, Appl
136	33.5	34.9	252	7	US-11-054-515-1236	Sequence 1236, Ap	209	32	33.3	72	6	US-10-986-501-162	Sequence 162, App
137	33.5	34.9	378	7	US-11-152-892-7	Sequence 7, Appli	210	32	33.3	95	7	US-11-054-669-68	Sequence 68, Appl
138	33.5	34.9	943	6	US-10-821-234-1012	Sequence 1012, Ap	211	32	33.3	95	7	US-11-084-554-133	Sequence 133, App
139	33	34.4	21	6	US-10-939-890-569	Sequence 569, App	212	32	33.3	95	7	US-11-061-848-45	Sequence 45, Appl
140	33	34.4	51	6	US-10-523-038-14	Sequence 14, Appl	213	32	33.3	95	7	US-11-061-848-64	Sequence 64, Appl
141	33	34.4	51	6	US-10-523-038-18	Sequence 18, Appl	214	32	33.3	95	7	US-11-004-590-75	Sequence 75, Appl
142	33	34.4	94	6	US-10-999-866-23	Sequence 23, Appl	215	32	33.3	95	7	US-11-136-250-133	Sequence 133, App
143	33	34.4	94	7	US-11-061-821-23	Sequence 23, Appl	216	32	33.3	100	7	US-11-053-076-125	Sequence 125, App
144	33	34.4	99	7	US-11-084-554-229	Sequence 229, App	217	32	33.3	104	7	US-11-084-554-226	Sequence 226, App
145	33	34.4	106	7	US-11-136-250-229	Sequence 229, App	218	32	33.3	104	7	US-11-136-250-226	Sequence 226, App
146	33	34.4	106	7	US-11-041-471-4	Sequence 4, Appli	219	32	33.3	106	7	US-11-174-186-5	Sequence 5, Appli
147	33	34.4	108	6	US-10-925-366A-147	Sequence 147, App	220	32	33.3	106	7	US-11-096-074-4	Sequence 4, Appli
148	33	34.4	109	6	US-10-771-257-101	Sequence 101, App	221	32	33.3	106	7	US-11-095-822-4	Sequence 4, Appli
149	33	34.4	109	6	US-10-771-257-105	Sequence 105, App	222	32	33.3	106	7	US-11-169-171-4	Sequence 4, Appli
150	33	34.4	109	6	US-10-771-257-107	Sequence 107, App	223	32	33.3	107	7	US-11-144-248-14	Sequence 14, Appl
151	33	34.4	109	6	US-11-051-568-21	Sequence 21, Appl	224	32	33.3	107	7	US-11-107-028-13	Sequence 13, Appl
152	33	34.4	175	6	US-10-967-527A-7	Sequence 7, Appli	225	32	33.3	107	7	US-11-107-028-14	Sequence 14, Appl
153	33	34.4	191	6	US-10-995-561-824	Sequence 824, App	226	32	33.3	107	7	US-11-107-028-15	Sequence 15, Appl
154	33	34.4	236	6	US-10-995-561-823	Sequence 823, App	227	32	33.3	107	7	US-11-144-222-14	Sequence 14, Appl
155	33	34.4	236	6	US-10-523-038-5	Sequence 5, Appli	228	32	33.3	107	7	US-11-182-343-14	Sequence 14, Appl
156	33	34.4	237	6	US-10-523-038-45	Sequence 45, Appl	229	32	33.3	108	6	US-10-925-366A-97	Sequence 97, Appl
157	33	34.4	272	6	US-10-986-501-120	Sequence 120, App	230	32	33.3	108	6	US-10-925-366A-146	Sequence 146, App
158	33	34.4	328	7	US-11-156-084-275	Sequence 275, App	231	32	33.3	108	6	US-10-925-366A-162	Sequence 162, App
159	33	34.4	340	7	US-11-092-353-1	Sequence 1, Appli	232	32	33.3	108	7	US-11-054-669-111	Sequence 111, App
160	33	34.4	361	7	US-11-186-284-173	Sequence 173, App	233	32	33.3	109	6	US-11-097-812-39	Sequence 39, Appl
161	33	34.4	418	7	US-11-082-389-50	Sequence 50, Appl	234	32	33.3	109	6	US-10-771-257-112	Sequence 112, App
162	33	34.4	422	7	US-11-055-822-560	Sequence 560, App	235	32	33.3	114	7	US-11-186-284-47	Sequence 47, Appl
163	33	34.4	422	7	US-11-055-822-562	Sequence 562, App	236	32	33.3	114	7	US-11-054-281-6	Sequence 6, Appli
164	33	34.4	460	7	US-11-182-016-26	Sequence 26, Appl	237	32	33.3	114	7	US-11-054-281-51	Sequence 51, Appl
165	33	34.4	463	7	US-11-069-642-107	Sequence 107, App	238	32	33.3	114	7	US-11-054-281-53	Sequence 53, Appl
166	33	34.4	472	6	US-10-650-326B-12	Sequence 12, Appl	239	32	33.3	124	7	US-11-096-074-8	Sequence 8, Appli
167	33	34.4	472	7	US-11-092-353-3	Sequence 3, Appli	240	32	33.3	124	7	US-11-095-822-8	Sequence 8, Appli
168	33	34.4	520	6	US-10-131-826A-144	Sequence 144, App	241	32	33.3	128	6	US-10-467-657-956	Sequence 956, App
169	33	34.4	537	7	US-11-129-442-47	Sequence 47, Appl	242	32	33.3	128	7	US-11-096-074-20	Sequence 20, Appl
170	33	34.4	589	7	US-11-183-136-42	Sequence 42, Appl	243	32	33.3	128	7	US-11-096-074-22	Sequence 22, Appl
171	33	34.4	617	7	US-11-143-980-35	Sequence 35, Appl	244	32	33.3	128	7	US-11-096-074-24	Sequence 24, Appl



245	32	33.3	128	7	US-11-095-822-20	Sequence 20, Appl	318	32	33.3	1398	6	US-10-453-372-872	Sequence 872, App
246	32	33.3	128	7	US-11-095-822-22	Sequence 22, Appl	319	32	33.3	1403	6	US-10-055-877-52	Sequence 52, Appl
247	32	33.3	128	7	US-11-095-822-24	Sequence 24, Appl	320	32	33.3	1403	6	US-10-453-372-878	Sequence 878, App
248	32	33.3	165	7	US-11-176-830-143	Sequence 143, App	321	32	33.3	1404	6	US-10-055-877-44	Sequence 44, Appl
249	32	33.3	165	7	US-11-196-067-143	Sequence 143, App	322	32	33.3	1404	6	US-10-453-372-870	Sequence 870, App
250	32	33.3	185	7	US-11-125-402-1	Sequence 1, Appl	323	32	33.3	1418	6	US-10-453-372-864	Sequence 864, App
251	32	33.3	212	6	US-10-510-101-71	Sequence 71, Appl	324	32	33.3	1450	6	US-10-055-877-48	Sequence 48, Appl
252	32	33.3	235	7	US-11-086-289-16	Sequence 16, Appl	325	32	33.3	1450	6	US-10-453-372-874	Sequence 874, App
253	32	33.3	237	6	US-10-073-301A-9	Sequence 9, Appl	326	32	33.3	1547	6	US-10-453-372-886	Sequence 886, App
254	32	33.3	237	6	US-11-054-669-109	Sequence 109, App	327	32	33.3	1577	6	US-10-055-877-54	Sequence 54, Appl
255	32	33.3	237	7	US-11-203-137-9	Sequence 9, Appl	328	32	33.3	1577	6	US-10-453-372-882	Sequence 882, App
256	32	33.3	248	7	US-11-054-515-1104	Sequence 1104, App	329	32	33.3	1577	6	US-10-453-372-884	Sequence 884, App
257	32	33.3	249	6	US-10-454-437-344	Sequence 344, App	330	32	33.3	1594	6	US-10-453-372-860	Sequence 860, App
258	32	33.3	252	6	US-10-763-712A-114	Sequence 114, App	331	32	33.3	1620	6	US-10-453-372-868	Sequence 868, App
259	32	33.3	254	7	US-11-054-515-1466	Sequence 1466, App	332	32	33.3	1653	6	US-10-453-372-866	Sequence 866, App
260	32	33.3	275	6	US-10-467-657-3548	Sequence 3548, App	333	32	33.3	1717	7	US-11-182-016-20	Sequence 20, Appl
261	32	33.3	275	6	US-10-467-657-5898	Sequence 5898, App	334	32	33.3	1857	6	US-10-055-877-252	Sequence 252, App
262	32	33.3	289	6	US-10-510-386-148	Sequence 148, App	335	32	33.3	2084	6	US-10-055-877-73	Sequence 73, Appl
263	32	33.3	293	7	US-11-055-822-1048	Sequence 1048, App	336	32	33.3	2109	6	US-10-055-877-251	Sequence 251, App
264	32	33.3	298	6	US-10-510-386-64	Sequence 64, Appl	337	32	33.3	2326	7	US-11-126-313-37	Sequence 37, Appl
265	32	33.3	305	7	US-11-024-959-266	Sequence 266, App	338	32	33.3	2432	6	US-10-821-234-899	Sequence 899, App
266	32	33.3	309	6	US-10-858-730-213	Sequence 213, App	339	32	33.3	2630	7	US-11-186-731-2	Sequence 2, Appl
267	32	33.3	311	7	US-11-082-389-168	Sequence 168, App	340	32	33.3	7968	7	US-11-186-731-5	Sequence 5, Appl
268	32	33.3	311	7	US-11-179-411-18	Sequence 18, Appl	341	31.5	32.8	165	6	US-10-821-234-1539	Sequence 1539, App
269	32	33.3	311	7	US-11-175-766-18	Sequence 18, Appl	342	31.5	32.8	174	6	US-10-519-390-4	Sequence 4, Appl
270	32	33.3	324	6	US-10-821-234-1471	Sequence 1471, App	343	31.5	32.8	174	7	US-11-033-365-144	Sequence 144, App
271	32	33.3	340	7	US-11-143-980-43	Sequence 43, Appl	344	31.5	32.8	174	7	US-11-113-581-82	Sequence 82, Appl
272	32	33.3	349	7	US-11-130-821-2	Sequence 2, Appl	345	31.5	32.8	174	7	US-11-052-472A-2	Sequence 2, Appl
273	32	33.3	354	6	US-10-878-556A-164	Sequence 164, App	346	31.5	32.8	175	7	US-11-033-365-143	Sequence 143, App
274	32	33.3	367	7	US-11-152-366-35	Sequence 35, Appl	347	31.5	32.8	175	7	US-11-033-365-147	Sequence 147, App
275	32	33.3	368	6	US-10-793-626-742	Sequence 742, App	348	31.5	32.8	175	7	US-11-033-365-151	Sequence 151, App
276	32	33.3	375	7	US-11-132-864-34	Sequence 34, Appl	349	31.5	32.8	175	7	US-11-033-365-157	Sequence 157, App
277	32	33.3	379	6	US-10-131-826A-328	Sequence 328, App	350	31.5	32.8	175	7	US-11-033-365-158	Sequence 158, App
278	32	33.3	388	7	US-11-130-821-1	Sequence 1, Appl	351	31.5	32.8	175	7	US-11-033-365-195	Sequence 195, App
279	32	33.3	393	7	US-11-077-712-4	Sequence 4, Appl	352	31.5	32.8	175	7	US-11-113-581-2	Sequence 2, Appl
280	32	33.3	393	7	US-11-109-831-25	Sequence 25, Appl	353	31.5	32.8	175	7	US-11-113-581-67	Sequence 67, Appl
281	32	33.3	399	6	US-10-926-709-17	Sequence 17, Appl	354	31.5	32.8	175	7	US-11-113-581-68	Sequence 68, Appl
282	32	33.3	459	7	US-11-186-284-12	Sequence 12, Appl	355	31.5	32.8	175	7	US-11-113-581-69	Sequence 69, Appl
283	32	33.3	582	7	US-11-090-439-58	Sequence 58, Appl	356	31.5	32.8	175	7	US-11-113-581-70	Sequence 70, Appl
284	32	33.3	582	7	US-11-169-041-130	Sequence 130, App	357	31.5	32.8	175	7	US-11-113-581-71	Sequence 71, Appl
285	32	33.3	714	7	US-11-121-419-2	Sequence 2, Appl	358	31.5	32.8	175	7	US-11-113-581-72	Sequence 72, Appl
286	32	33.3	751	7	US-11-114-906-8	Sequence 8, Appl	359	31.5	32.8	175	7	US-11-113-581-73	Sequence 73, Appl
287	32	33.3	764	7	US-11-114-906-6	Sequence 6, Appl	360	31.5	32.8	175	7	US-11-113-581-74	Sequence 74, Appl
288	32	33.3	765	7	US-11-024-959-476	Sequence 476, App	361	31.5	32.8	175	7	US-11-113-581-75	Sequence 75, Appl
289	32	33.3	776	7	US-11-114-906-24	Sequence 24, Appl	362	31.5	32.8	175	7	US-11-113-581-76	Sequence 76, Appl
290	32	33.3	783	7	US-11-186-284-59	Sequence 59, Appl	363	31.5	32.8	175	7	US-11-113-581-77	Sequence 77, Appl
291	32	33.3	789	7	US-11-114-906-22	Sequence 22, Appl	364	31.5	32.8	175	7	US-11-113-581-78	Sequence 78, Appl
292	32	33.3	838	7	US-11-114-906-40	Sequence 40, Appl	365	31.5	32.8	175	7	US-11-113-581-79	Sequence 79, Appl
293	32	33.3	851	7	US-11-114-906-38	Sequence 38, Appl	366	31.5	32.8	175	7	US-11-113-581-80	Sequence 80, Appl
294	32	33.3	859	7	US-11-000-463-423	Sequence 423, App	367	31.5	32.8	175	7	US-11-113-581-81	Sequence 81, Appl
295	32	33.3	859	7	US-11-000-463-895	Sequence 895, App	368	31.5	32.8	175	7	US-11-113-581-83	Sequence 83, Appl
296	32	33.3	863	7	US-11-114-906-32	Sequence 32, Appl	369	31.5	32.8	175	7	US-11-113-581-84	Sequence 84, Appl
297	32	33.3	864	7	US-11-114-906-4	Sequence 4, Appl	370	31.5	32.8	175	7	US-11-113-581-85	Sequence 85, Appl
298	32	33.3	870	7	US-11-114-906-2	Sequence 2, Appl	371	31.5	32.8	175	7	US-11-113-581-86	Sequence 86, Appl
299	32	33.3	875	6	US-10-858-730-7	Sequence 7, Appl	372	31.5	32.8	175	7	US-11-113-581-87	Sequence 87, Appl
300	32	33.3	876	7	US-11-114-906-30	Sequence 30, Appl	373	31.5	32.8	175	7	US-11-113-581-88	Sequence 88, Appl
301	32	33.3	889	7	US-11-114-906-20	Sequence 20, Appl	374	31.5	32.8	175	7	US-11-113-581-89	Sequence 89, Appl
302	32	33.3	895	7	US-11-114-906-18	Sequence 18, Appl	375	31.5	32.8	175	7	US-11-113-581-91	Sequence 91, Appl
303	32	33.3	909	7	US-11-120-308-190	Sequence 190, App	376	31.5	32.8	175	7	US-11-113-581-92	Sequence 92, Appl
304	32	33.3	918	6	US-10-995-561-981	Sequence 981, App	377	31.5	32.8	175	7	US-11-113-581-93	Sequence 93, Appl
305	32	33.3	922	7	US-11-115-086-9	Sequence 9, Appl	378	31.5	32.8	175	7	US-11-113-581-94	Sequence 94, Appl
306	32	33.3	932	7	US-11-057-058-59	Sequence 59, App	379	31.5	32.8	175	7	US-11-113-581-95	Sequence 95, Appl
307	32	33.3	951	7	US-11-114-906-36	Sequence 36, Appl	380	31.5	32.8	175	7	US-11-113-581-96	Sequence 96, Appl
308	32	33.3	957	7	US-11-114-906-34	Sequence 34, Appl	381	31.5	32.8	175	7	US-11-113-581-97	Sequence 97, Appl
309	32	33.3	974	7	US-11-115-086-7	Sequence 7, Appl	382	31.5	32.8	175	7	US-11-113-581-98	Sequence 98, Appl
310	32	33.3	976	7	US-11-114-906-28	Sequence 28, Appl	383	31.5	32.8	175	7	US-11-113-581-99	Sequence 99, Appl
311	32	33.3	982	7	US-11-114-906-26	Sequence 26, Appl	384	31.5	32.8	175	7	US-11-113-581-100	Sequence 100, App
312	32	33.3	1123	7	US-11-037-243-77	Sequence 77, Appl	385	31.5	32.8	175	7	US-11-113-581-102	Sequence 102, App
313	32	33.3	1198	6	US-10-453-372-880	Sequence 880, App	386	31.5	32.8	175	7	US-11-113-581-103	Sequence 103, App
314	32	33.3	1251	7	US-11-043-889-22	Sequence 22, Appl	387	31.5	32.8	175	7	US-11-113-581-104	Sequence 104, App
315	32	33.3	1366	6	US-10-821-234-1431	Sequence 1431, App	388	31.5	32.8	175	7	US-11-113-581-105	Sequence 105, App
316	32	33.3	1366	7	US-11-186-284-31	Sequence 31, Appl	389	31.5	32.8	175	7	US-11-113-581-106	Sequence 106, App
317	32	33.3	1398	6	US-10-055-877-46	Sequence 46, Appl	390	31.5	32.8	175	7	US-11-113-581-107	Sequence 107, App

391	31.5	32.8	175	7	US-11-113-581-108	Sequence 108, App	464	31	32.3	95	7	US-11-054-669-90	Sequence 90, Appl
392	31.5	32.8	175	7	US-11-113-581-109	Sequence 109, App	465	31	32.3	95	7	US-11-084-554-119	Sequence 119, App
393	31.5	32.8	175	7	US-11-113-581-110	Sequence 110, App	466	31	32.3	95	7	US-11-084-554-127	Sequence 127, App
394	31.5	32.8	175	7	US-11-052-472A-1	Sequence 1, Appli	467	31	32.3	95	7	US-11-084-554-130	Sequence 130, App
395	31.5	32.8	175	7	US-11-052-472A-3	Sequence 3, Appli	468	31	32.3	95	7	US-11-084-554-131	Sequence 131, App
396	31.5	32.8	176	7	US-11-033-365-145	Sequence 145, App	469	31	32.3	95	7	US-11-084-554-138	Sequence 138, App
397	31.5	32.8	176	7	US-11-033-365-146	Sequence 146, App	470	31	32.3	95	7	US-11-061-821-24	Sequence 24, Appl
398	31.5	32.8	176	7	US-11-033-365-148	Sequence 148, App	471	31	32.3	95	7	US-11-061-848-32	Sequence 32, Appl
399	31.5	32.8	176	7	US-11-033-365-149	Sequence 149, App	472	31	32.3	95	7	US-11-061-848-50	Sequence 50, Appl
400	31.5	32.8	177	7	US-11-033-365-142	Sequence 142, App	473	31	32.3	95	7	US-11-061-848-51	Sequence 51, Appl
401	31.5	32.8	177	7	US-11-033-365-152	Sequence 152, App	474	31	32.3	95	7	US-11-061-848-52	Sequence 52, Appl
402	31.5	32.8	177	7	US-11-033-365-154	Sequence 154, App	475	31	32.3	95	7	US-11-061-848-65	Sequence 65, Appl
403	31.5	32.8	177	7	US-11-033-365-194	Sequence 194, App	476	31	32.3	95	7	US-11-061-848-66	Sequence 66, Appl
404	31.5	32.8	177	7	US-11-176-830-210	Sequence 210, App	477	31	32.3	95	7	US-11-061-848-67	Sequence 67, Appl
405	31.5	32.8	177	7	US-11-176-830-631	Sequence 631, App	478	31	32.3	95	7	US-11-061-848-69	Sequence 69, Appl
406	31.5	32.8	177	7	US-11-176-830-632	Sequence 632, App	479	31	32.3	95	7	US-11-004-590-88	Sequence 88, Appl
407	31.5	32.8	177	7	US-11-176-830-633	Sequence 633, App	480	31	32.3	95	7	US-11-004-590-89	Sequence 89, Appl
408	31.5	32.8	177	7	US-11-176-830-634	Sequence 634, App	481	31	32.3	95	7	US-11-004-590-90	Sequence 90, Appl
409	31.5	32.8	177	7	US-11-176-830-635	Sequence 635, App	482	31	32.3	95	7	US-11-004-590-91	Sequence 91, Appl
410	31.5	32.8	177	7	US-11-176-830-636	Sequence 636, App	483	31	32.3	95	7	US-11-004-590-95	Sequence 95, Appl
411	31.5	32.8	177	7	US-11-176-830-637	Sequence 637, App	484	31	32.3	95	7	US-11-136-250-119	Sequence 119, App
412	31.5	32.8	177	7	US-11-176-830-638	Sequence 638, App	485	31	32.3	95	7	US-11-136-250-127	Sequence 127, App
413	31.5	32.8	177	7	US-11-176-830-639	Sequence 639, App	486	31	32.3	95	7	US-11-136-250-130	Sequence 130, App
414	31.5	32.8	177	7	US-11-176-830-640	Sequence 640, App	487	31	32.3	95	7	US-11-136-250-131	Sequence 131, App
415	31.5	32.8	177	7	US-11-176-830-641	Sequence 641, App	488	31	32.3	95	7	US-11-136-250-138	Sequence 138, App
416	31.5	32.8	177	7	US-11-176-830-642	Sequence 642, App	489	31	32.3	96	7	US-11-144-248-38	Sequence 38, Appl
417	31.5	32.8	177	7	US-11-176-830-643	Sequence 643, App	490	31	32.3	96	7	US-11-054-669-82	Sequence 82, Appl
418	31.5	32.8	177	7	US-11-176-830-644	Sequence 644, App	491	31	32.3	96	7	US-11-054-669-88	Sequence 88, Appl
419	31.5	32.8	177	7	US-11-176-830-645	Sequence 645, App	492	31	32.3	96	7	US-11-084-554-114	Sequence 114, App
420	31.5	32.8	177	7	US-11-176-830-646	Sequence 646, App	493	31	32.3	96	7	US-11-084-554-135	Sequence 135, App
421	31.5	32.8	177	7	US-11-176-830-647	Sequence 647, App	494	31	32.3	96	7	US-11-144-222-38	Sequence 38, Appl
422	31.5	32.8	177	7	US-11-176-830-648	Sequence 648, App	495	31	32.3	96	7	US-11-051-453-82	Sequence 82, Appl
423	31.5	32.8	177	7	US-11-176-830-655	Sequence 655, App	496	31	32.3	96	7	US-11-128-900-87	Sequence 87, Appl
424	31.5	32.8	177	7	US-11-176-830-656	Sequence 656, App	497	31	32.3	96	7	US-11-004-590-86	Sequence 86, Appl
425	31.5	32.8	177	7	US-11-176-830-657	Sequence 657, App	498	31	32.3	96	7	US-11-004-590-87	Sequence 87, Appl
426	31.5	32.8	177	7	US-11-176-830-658	Sequence 658, App	499	31	32.3	96	7	US-11-004-590-92	Sequence 92, Appl
427	31.5	32.8	177	7	US-11-176-830-659	Sequence 659, App	500	31	32.3	96	7	US-11-136-250-114	Sequence 114, App
428	31.5	32.8	177	7	US-11-176-830-660	Sequence 660, App	501	31	32.3	96	7	US-11-136-250-135	Sequence 135, App
429	31.5	32.8	177	7	US-11-176-830-661	Sequence 661, App	502	31	32.3	96	7	US-11-182-343-38	Sequence 38, Appl
430	31.5	32.8	177	7	US-11-176-830-662	Sequence 662, App	503	31	32.3	97	7	US-11-084-554-122	Sequence 122, App
431	31.5	32.8	178	7	US-11-033-365-141	Sequence 141, App	504	31	32.3	97	7	US-11-136-250-122	Sequence 122, App
432	31.5	32.8	178	7	US-11-033-365-155	Sequence 155, App	505	31	32.3	98	7	US-11-084-554-230	Sequence 230, App
433	31.5	32.8	179	7	US-11-033-365-153	Sequence 153, App	506	31	32.3	98	7	US-11-136-250-230	Sequence 230, App
434	31.5	32.8	181	7	US-11-033-365-150	Sequence 150, App	507	31	32.3	100	7	US-11-144-248-10	Sequence 10, Appl
435	31.5	32.8	256	7	US-11-054-515-1027	Sequence 1027, Ap	508	31	32.3	100	7	US-11-144-222-10	Sequence 10, Appl
436	31.5	32.8	258	7	US-11-054-515-865	Sequence 865, App	509	31	32.3	100	7	US-11-182-343-10	Sequence 10, Appl
437	31.5	32.8	258	7	US-11-054-515-1029	Sequence 1029, Ap	510	31	32.3	104	7	US-11-084-554-222	Sequence 222, App
438	31.5	32.8	259	7	US-11-054-515-1277	Sequence 1277, Ap	511	31	32.3	104	7	US-11-136-250-222	Sequence 222, App
439	31.5	32.8	259	7	US-11-054-515-1481	Sequence 1481, Ap	512	31	32.3	106	6	US-10-467-657-4996	Sequence 4996, Ap
440	31.5	32.8	260	7	US-11-054-515-1036	Sequence 1036, Ap	513	31	32.3	106	7	US-11-174-186-1	Sequence 1, Appli
441	31.5	32.8	279	6	US-11-078-556A-157	Sequence 157, App	514	31	32.3	106	7	US-11-174-186-8	Sequence 8, Appli
442	31.5	32.8	338	7	US-11-055-822-584	Sequence 584, App	515	31	32.3	106	7	US-11-174-186-9	Sequence 9, Appli
443	31.5	32.8	508	6	US-11-078-556A-154	Sequence 154, App	516	31	32.3	106	7	US-11-174-186-15	Sequence 15, Appl
444	31.5	32.8	818	7	US-11-037-243-111	Sequence 111, App	517	31	32.3	106	7	US-11-155-775-8	Sequence 8, Appli
445	31.5	32.8	908	6	US-10-821-234-1267	Sequence 1267, Ap	518	31	32.3	106	7	US-11-155-775-32	Sequence 32, Appl
446	31.5	32.8	1302	7	US-11-004-057-6	Sequence 6, Appli	519	31	32.3	106	7	US-11-149-943-54	Sequence 54, Appl
447	31.5	32.8	1493	7	US-11-004-057-4	Sequence 4, Appli	520	31	32.3	107	6	US-10-942-440-62	Sequence 62, Appl
448	31.5	32.8	1493	7	US-11-004-057-21	Sequence 21, Appl	521	31	32.3	107	7	US-11-040-159-17	Sequence 17, Appl
449	31	32.3	46	6	US-10-986-501-109	Sequence 109, App	522	31	32.3	107	7	US-11-093-274-24	Sequence 24, Appl
450	31	32.3	60	7	US-11-091-668-15	Sequence 15, Appl	523	31	32.3	107	7	US-11-069-834-8	Sequence 8, Appli
451	31	32.3	79	6	US-10-999-866-14	Sequence 14, Appl	524	31	32.3	108	6	US-10-925-366A-150	Sequence 150, App
452	31	32.3	79	7	US-11-061-821-14	Sequence 14, Appl	525	31	32.3	108	6	US-10-925-366A-193	Sequence 193, App
453	31	32.3	91	6	US-10-999-866-12	Sequence 12, Appl	526	31	32.3	108	6	US-10-925-366A-205	Sequence 205, App
454	31	32.3	91	7	US-11-144-248-22	Sequence 22, Appl	527	31	32.3	108	6	US-10-925-366A-207	Sequence 207, App
455	31	32.3	91	7	US-11-061-821-12	Sequence 12, Appl	528	31	32.3	108	6	US-10-925-366A-332	Sequence 332, App
456	31	32.3	91	7	US-11-144-221-22	Sequence 22, Appl	529	31	32.3	108	6	US-10-771-257-33	Sequence 33, Appl
457	31	32.3	91	7	US-11-182-343-22	Sequence 22, Appl	530	31	32.3	108	6	US-10-850-635-4	Sequence 4, Appli
458	31	32.3	94	7	US-11-093-274-33	Sequence 33, Appl	531	31	32.3	108	6	US-10-850-635-6	Sequence 6, Appli
459	31	32.3	95	6	US-10-999-866-24	Sequence 24, Appl	532	31	32.3	108	6	US-10-886-383-2	Sequence 2, Appli
460	31	32.3	95	7	US-11-054-669-84	Sequence 84, Appl	533	31	32.3	108	6	US-10-886-383-4	Sequence 4, Appli
461	31	32.3	95	7	US-11-054-669-85	Sequence 85, Appl	534	31	32.3	108	6	US-10-982-440-32	Sequence 32, Appl
462	31	32.3	95	7	US-11-054-669-86	Sequence 86, Appl	535	31	32.3	108	6	US-10-982-440-36	Sequence 36, Appl
463	31	32.3	95	7	US-11-054-669-87	Sequence 87, Appl	536	31	32.3	108	7	US-11-040-159-15	Sequence 15, Appl



683	31	32.3	251	7	US-11-054-515-83	Sequence 83, App1
684	31	32.3	251	7	US-11-054-515-84	Sequence 84, App1
685	31	32.3	251	7	US-11-054-515-85	Sequence 85, App1
686	31	32.3	251	7	US-11-054-515-86	Sequence 86, App1
687	31	32.3	251	7	US-11-054-515-87	Sequence 87, App1
688	31	32.3	251	7	US-11-054-515-88	Sequence 88, App1
689	31	32.3	251	7	US-11-054-515-89	Sequence 89, App1
690	31	32.3	251	7	US-11-054-515-90	Sequence 90, App1
691	31	32.3	251	7	US-11-054-515-91	Sequence 91, App1
692	31	32.3	251	7	US-11-054-515-92	Sequence 92, App1
693	31	32.3	251	7	US-11-054-515-93	Sequence 93, App1
694	31	32.3	251	7	US-11-054-515-94	Sequence 94, App1
695	31	32.3	251	7	US-11-054-515-95	Sequence 95, App1
696	31	32.3	251	7	US-11-054-515-96	Sequence 96, App1
697	31	32.3	251	7	US-11-054-515-97	Sequence 97, App1
698	31	32.3	251	7	US-11-054-515-98	Sequence 98, App1
699	31	32.3	251	7	US-11-054-515-99	Sequence 99, App1
700	31	32.3	251	7	US-11-054-515-101	Sequence 101, App
701	31	32.3	251	7	US-11-054-515-102	Sequence 102, App
702	31	32.3	251	7	US-11-054-515-103	Sequence 103, App
703	31	32.3	251	7	US-11-054-515-104	Sequence 104, App
704	31	32.3	251	7	US-11-054-515-105	Sequence 105, App
705	31	32.3	251	7	US-11-054-515-106	Sequence 106, App
706	31	32.3	251	7	US-11-054-515-108	Sequence 108, App
707	31	32.3	251	7	US-11-054-515-109	Sequence 109, App
708	31	32.3	251	7	US-11-054-515-110	Sequence 110, App
709	31	32.3	251	7	US-11-054-515-112	Sequence 112, App
710	31	32.3	251	7	US-11-054-515-113	Sequence 113, App
711	31	32.3	251	7	US-11-054-515-114	Sequence 114, App
712	31	32.3	251	7	US-11-054-515-115	Sequence 115, App
713	31	32.3	251	7	US-11-054-515-116	Sequence 116, App
714	31	32.3	251	7	US-11-054-515-118	Sequence 118, App
715	31	32.3	251	7	US-11-054-515-119	Sequence 119, App
716	31	32.3	251	7	US-11-054-515-120	Sequence 120, App
717	31	32.3	251	7	US-11-054-515-121	Sequence 121, App
718	31	32.3	251	7	US-11-054-515-122	Sequence 122, App
719	31	32.3	251	7	US-11-054-515-123	Sequence 123, App
720	31	32.3	251	7	US-11-054-515-124	Sequence 124, App
721	31	32.3	251	7	US-11-054-515-125	Sequence 125, App
722	31	32.3	251	7	US-11-054-515-127	Sequence 127, App
723	31	32.3	251	7	US-11-054-515-128	Sequence 128, App
724	31	32.3	251	7	US-11-054-515-129	Sequence 129, App
725	31	32.3	251	7	US-11-054-515-130	Sequence 130, App
726	31	32.3	251	7	US-11-054-515-131	Sequence 131, App
727	31	32.3	251	7	US-11-054-515-132	Sequence 132, App
728	31	32.3	251	7	US-11-054-515-133	Sequence 133, App
729	31	32.3	251	7	US-11-054-515-134	Sequence 134, App
730	31	32.3	251	7	US-11-054-515-135	Sequence 135, App
731	31	32.3	251	7	US-11-054-515-136	Sequence 136, App
732	31	32.3	251	7	US-11-054-515-144	Sequence 144, App
733	31	32.3	251	7	US-11-054-515-137	Sequence 137, App
734	31	32.3	251	7	US-11-054-515-141	Sequence 141, App
735	31	32.3	251	7	US-11-054-515-142	Sequence 142, App
736	31	32.3	251	7	US-11-054-515-143	Sequence 143, App
737	31	32.3	251	7	US-11-054-515-145	

829	31	32.3	251	7	US-11-054-515-245	Sequence 245, App	902	31	32.3	276	6	US-10-873-528-134	Sequence 134, App
830	31	32.3	251	7	US-11-054-515-246	Sequence 246, App	903	31	32.3	282	6	US-10-467-657-5742	Sequence 5742, App
831	31	32.3	251	7	US-11-054-515-247	Sequence 247, App	904	31	32.3	288	7	US-11-149-462-8	Sequence 8, Appli
832	31	32.3	251	7	US-11-054-515-248	Sequence 248, App	905	31	32.3	296	6	US-10-965-972-8	Sequence 355, App
833	31	32.3	251	7	US-11-054-515-249	Sequence 249, App	906	31	32.3	302	7	US-11-156-084-355	Sequence 332, App
834	31	32.3	251	7	US-11-054-515-250	Sequence 250, App	907	31	32.3	307	7	US-11-000-463-332	Sequence 334, App
835	31	32.3	251	7	US-11-054-515-251	Sequence 251, App	908	31	32.3	312	7	US-11-000-463-334	Sequence 1192, App
836	31	32.3	251	7	US-11-054-515-252	Sequence 252, App	909	31	32.3	313	7	US-11-156-084-337	Sequence 64, Appli
837	31	32.3	251	7	US-11-054-515-253	Sequence 253, App	910	31	32.3	313	7	US-11-156-084-338	Sequence 6, Appli
838	31	32.3	251	7	US-11-054-515-254	Sequence 254, App	911	31	32.3	313	7	US-11-156-084-338	Sequence 14, Appli
839	31	32.3	251	7	US-11-054-515-255	Sequence 255, App	912	31	32.3	321	6	US-10-821-234-1192	Sequence 13, Appli
840	31	32.3	251	7	US-11-054-515-256	Sequence 256, App	913	31	32.3	325	6	US-10-063-703-64	Sequence 6, Appli
841	31	32.3	251	7	US-11-054-515-257	Sequence 257, App	914	31	32.3	325	7	US-11-102-240-64	Sequence 6, Appli
842	31	32.3	251	7	US-11-054-515-258	Sequence 258, App	915	31	32.3	335	7	US-11-018-018-6	Sequence 6, Appli
843	31	32.3	251	7	US-11-054-515-259	Sequence 259, App	916	31	32.3	335	7	US-11-047-757-6	Sequence 14, Appli
844	31	32.3	251	7	US-11-054-515-260	Sequence 260, App	917	31	32.3	335	7	US-11-182-946-14	Sequence 13, Appli
845	31	32.3	251	7	US-11-054-515-261	Sequence 261, App	918	31	32.3	335	7	US-11-166-892-13	Sequence 6, Appli
846	31	32.3	251	7	US-11-054-515-262	Sequence 262, App	919	31	32.3	335	7	US-11-048-490-6	Sequence 13, Appli
847	31	32.3	251	7	US-11-054-515-263	Sequence 263, App	920	31	32.3	335	7	US-11-048-490-6	Sequence 2544, App
848	31	32.3	251	7	US-11-054-515-264	Sequence 264, App	921	31	32.3	335	7	US-11-048-490-6	Sequence 270, App
849	31	32.3	251	7	US-11-054-515-265	Sequence 265, App	922	31	32.3	335	7	US-11-048-490-6	Sequence 20, Appli
850	31	32.3	251	7	US-11-054-515-266	Sequence 266, App	923	31	32.3	335	7	US-11-048-490-6	Sequence 804, App
851	31	32.3	251	7	US-11-054-515-267	Sequence 267, App	924	31	32.3	335	7	US-11-000-463-805	Sequence 805, App
852	31	32.3	251	7	US-11-054-515-268	Sequence 268, App	925	31	32.3	335	7	US-11-000-463-806	Sequence 806, App
853	31	32.3	251	7	US-11-054-515-269	Sequence 269, App	926	31	32.3	335	7	US-11-000-463-807	Sequence 807, App
854	31	32.3	251	7	US-11-054-515-270	Sequence 270, App	927	31	32.3	335	7	US-11-000-463-807	Sequence 5, Appli
855	31	32.3	251	7	US-11-054-515-271	Sequence 271, App	928	31	32.3	335	7	US-11-000-463-807	Sequence 7, Appli
856	31	32.3	251	7	US-11-054-515-272	Sequence 272, App	929	31	32.3	335	7	US-11-000-463-807	Sequence 322, App
857	31	32.3	251	7	US-11-054-515-273	Sequence 273, App	930	31	32.3	335	7	US-11-000-463-807	Sequence 7, Appli
858	31	32.3	251	7	US-11-054-515-274	Sequence 274, App	931	31	32.3	335	7	US-11-000-463-807	Sequence 7, Appli
859	31	32.3	251	7	US-11-054-515-275	Sequence 275, App	932	31	32.3	335	7	US-11-000-463-807	Sequence 5, Appli
860	31	32.3	251	7	US-11-054-515-276	Sequence 276, App	933	31	32.3	335	7	US-11-000-463-807	Sequence 276, App
861	31	32.3	251	7	US-11-054-515-277	Sequence 277, App	934	31	32.3	335	7	US-11-000-463-807	Sequence 5, Appli
862	31	32.3	251	7	US-11-054-515-278	Sequence 278, App	935	31	32.3	335	7	US-11-000-463-807	Sequence 43, Appli
863	31	32.3	251	7	US-11-054-515-279	Sequence 279, App	936	31	32.3	335	7	US-11-000-463-807	Sequence 3, Appli
864	31	32.3	251	7	US-11-054-515-280	Sequence 280, App	937	31	32.3	335	7	US-11-000-463-807	Sequence 274, App
865	31	32.3	251	7	US-11-054-515-281	Sequence 281, App	938	31	32.3	335	7	US-11-000-463-807	Sequence 341, App
866	31	32.3	251	7	US-11-054-515-282	Sequence 282, App	939	31	32.3	335	7	US-11-000-463-807	Sequence 6, Appli
867	31	32.3	251	7	US-11-054-515-283	Sequence 283, App	940	31	32.3	335	7	US-11-000-463-807	Sequence 5, Appli
868	31	32.3	251	7	US-11-054-515-284	Sequence 284, App	941	31	32.3	335	7	US-11-000-463-807	Sequence 6, Appli
869	31	32.3	251	7	US-11-054-515-285	Sequence 285, App	942	31	32.3	335	7	US-11-000-463-807	Sequence 14, Appli
870	31	32.3	251	7	US-11-054-515-286	Sequence 286, App	943	31	32.3	335	7	US-11-000-463-807	Sequence 950, App
871	31	32.3	251	7	US-11-054-515-287	Sequence 287, App	944	31	32.3	335	7	US-11-000-463-807	Sequence 3, Appli
872	31	32.3	251	7	US-11-054-515-288	Sequence 288, App	945	31	32.3	335	7	US-11-000-463-807	Sequence 2, Appli
873	31	32.3	251	7	US-11-054-515-289	Sequence 289, App	946	31	32.3	335	7	US-11-000-463-807	Sequence 36, Appli
874	31	32.3	251	7	US-11-054-515-290	Sequence 290, App	947	31	32.3	335	7	US-11-000-463-807	Sequence 6, Appli
875	31	32.3	251	7	US-11-054-515-291	Sequence 291, App	948	31	32.3	335	7	US-11-000-463-807	Sequence 22, Appli
876	31	32.3	251	7	US-11-054-515-292	Sequence 292, App	949	31	32.3	335	7	US-11-000-463-807	Sequence 26, Appli
877	31	32.3	251	7	US-11-054-515-293	Sequence 293, App	950	31	32.3	335	7	US-11-000-463-807	Sequence 28, Appli
878	31	32.3	251	7	US-11-054-515-294	Sequence 294, App	951	31	32.3	335	7	US-11-000-463-807	Sequence 30, Appli
879	31	32.3	251	7	US-11-054-515-295	Sequence 295, App	952	31	32.3	335	7	US-11-000-463-807	Sequence 10, Appli
880	31	32.3	251	7	US-11-054-515-296	Sequence 296, App	953	31	32.3	335	7	US-11-000-463-807	Sequence 34, Appli
881	31	32.3	251	7	US-11-054-515-297	Sequence 297, App	954	31	32.3	335	7	US-11-000-463-807	Sequence 14, Appli
882	31	32.3	251	7	US-11-054-515-298	Sequence 298, App	955	31	32.3	335	7	US-11-000-463-807	Sequence 950, App
883	31	32.3	251	7	US-11-054-515-299	Sequence 299, App	956	31	32.3	335	7	US-11-000-463-807	Sequence 3, Appli
884	31	32.3	251	7	US-11-054-515-300	Sequence 300, App	957	31	32.3	335	7	US-11-000-463-807	Sequence 2, Appli
885	31	32.3	251	7	US-11-054-515-301	Sequence 301, App	958	31	32.3	335	7	US-11-000-463-807	Sequence 36, Appli
886	31	32.3	251	7	US-11-054-515-302	Sequence 302, App	959	31	32.3	335	7	US-11-000-463-807	Sequence 6, Appli
887	31	32.3	251	7	US-11-054-515-303	Sequence 303, App	960	31	32.3	335	7	US-11-000-463-807	Sequence 22, Appli
888	31	32.3	251	7	US-11-054-515-304	Sequence 304, App	961	31	32.3	335	7	US-11-000-463-807	Sequence 26, Appli
889	31	32.3	251	7	US-11-054-515-305	Sequence 305, App	962	31	32.3	335	7	US-11-000-463-807	Sequence 28, Appli
890	31	32.3	251	7	US-11-054-515-306	Sequence 306, App	963	31	32.3	335	7	US-11-000-463-807	Sequence 30, Appli
891	31	32.3	251	7	US-11-054-515-307	Sequence 307, App	964	31	32.3	335	7	US-11-000-463-807	Sequence 10, Appli
892	31	32.3	251	7	US-11-054-515-308	Sequence 308, App	965	31	32.3	335	7	US-11-000-463-807	Sequence 34, Appli
893	31	32.3	251	7	US-11-054-515-309	Sequence 309, App	966	31	32.3	335	7	US-11-000-463-807	Sequence 14, Appli
894	31	32.3	251	7	US-11-054-515-310	Sequence 310, App	967	31	32.3	335	7	US-11-000-463-807	Sequence 950, App
895	31	32.3	251	7	US-11-054-515-311	Sequence 311, App	968	31	32.3	335	7	US-11-000-463-807	Sequence 3, Appli
896	31	32.3	251	7	US-11-054-515-312	Sequence 312, App	969	31	32.3	335	7	US-11-000-463-807	Sequence 2, Appli
897	31	32.3	251	7	US-11-054-515-313	Sequence 313, App	970	31	32.3	335	7	US-11-000-463-807	Sequence 36, Appli
898	31	32.3	251	7	US-11-054-515-314	Sequence 314, App	971	31	32.3	335	7	US-11-000-463-807	Sequence 6, Appli
899	31	32.3	251	7	US-11-054-515-315	Sequence 315, App	972	31	32.3	335	7	US-11-000-463-807	Sequence 22, Appli
900	31	32.3	251	7	US-11-054-515-316	Sequence 316, App	973	31	32.3	335	7	US-11-000-463-807	Sequence 26, Appli
901	31	32.3	251	7	US-11-054-515-317	Sequence 317, App	974	31	32.3	335	7	US-11-000-463-807	Sequence 28, Appli
					Sequence 318, App								
					Sequence 319, App								
					Sequence 320, App								
					Sequence 1049, App								
					Sequence 1219, App								
					Sequence 1534, App								
					Sequence 1499, App								
					Sequence 4, Appli								
					Sequence 27, Appli								

975 31 32.3 708 6 US-10-821-234-917 Sequence 917, App  
976 31 32.3 708 7 US-11-150-533-65 Sequence 65, Appl  
977 31 32.3 745 7 US-11-081-018-4 Sequence 4, Appli  
978 31 32.3 770 6 US-10-821-234-1269 Sequence 1269, Ap  
979 31 32.3 775 6 US-10-795-201-1 Sequence 1, Appli  
980 31 32.3 849 6 US-10-467-962B-53 Sequence 53, Appl  
981 31 32.3 863 7 US-11-169-041-167 Sequence 167, App  
982 31 32.3 1085 6 US-10-523-477-13 Sequence 13, Appl  
983 31 32.3 1206 6 US-10-995-561-709 Sequence 709, App  
984 31 32.3 1258 7 US-11-121-438-16 Sequence 16, Appl  
985 31 32.3 1307 6 US-10-995-561-711 Sequence 711, App  
986 31 32.3 1323 6 US-10-517-939-312 Sequence 312, App  
987 31 32.3 1360 7 US-11-188-743-22 Sequence 22, Appl  
988 31 32.3 1360 7 US-11-183-294-26 Sequence 26, Appl  
989 31 32.3 1579 7 US-11-052-554-9 Sequence 9, Appli  
990 31 32.3 1613 7 US-11-108-528-86 Sequence 86, Appl  
991 31 32.3 1874 6 US-10-821-234-1182 Sequence 1182, Ap  
992 31 32.3 1985 6 US-10-495-083-7 Sequence 7, Appli  
993 31 32.3 2161 7 US-11-126-313-31 Sequence 31, Appl  
994 31 32.3 2256 7 US-11-144-368-4 Sequence 4, Appli  
995 31 32.3 2455 7 US-11-186-999-4 Sequence 4, Appli  
996 31 32.3 2455 7 US-11-186-999-14 Sequence 14, Appl  
997 31 32.3 2455 7 US-11-186-999-16 Sequence 16, Appl  
998 31 32.3 2458 7 US-11-186-999-6 Sequence 6, Appli  
999 31 32.3 2458 7 US-11-186-999-11 Sequence 11, Appl  
1000 31 32.3 2458 7 US-11-186-999-13 Sequence 13, Appl

## ALIGNMENTS

RESULT 1  
US-10-945-853-3  
; Sequence 3, Application US/10945853  
; Publication No. US20050255117A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc.  
; APPLICANT: Sanicola-Nadel, Michele  
; APPLICANT: Adkins, Heather  
; APPLICANT: Miklasz, Steven Donald  
; APPLICANT: Rayhorn, Paul  
; APPLICANT: Schiffer, Susan Gail  
; APPLICANT: Williams, Kevin  
; TITLE OF INVENTION: Crypto-Specific Antibodies  
; FILE REFERENCE: BGN117CPCCN  
; CURRENT FILING DATE: 2004-09-20  
; PRIOR FILING DATE: 2004-09-20  
; PRIOR APPLICATION NUMBER: US/10/945,853  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: PCT/US02/31462  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: PCT/US02/11950  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/286,782  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: 60/293,020  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/301,091  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/367,002  
; PRIOR FILING DATE: 2002-03-22  
; NUMBER OF SEQ ID NOS: 9  
; SEQ ID NO 3  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-945-853-3  
Query Match 100.0%; Score 96; DB 6; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.2e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 FRDSDIWPQEEPAIRPR 17  
Db 1 FRDSDIWPQEEPAIRPR 17  
|||||

RESULT 2  
US-10-945-853-1  
; Sequence 1, Application US/10945853  
; Publication No. US20050255117A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc.  
; APPLICANT: Sanicola-Nadel, Michele  
; APPLICANT: Adkins, Heather  
; APPLICANT: Miklasz, Steven Donald  
; APPLICANT: Rayhorn, Paul  
; APPLICANT: Schiffer, Susan Gail  
; APPLICANT: Williams, Kevin  
; TITLE OF INVENTION: Crypto-Specific Antibodies  
; FILE REFERENCE: BGN117CPCCN  
; CURRENT FILING DATE: 2004-09-20  
; PRIOR FILING DATE: 2004-09-20  
; PRIOR APPLICATION NUMBER: PCT/US02/31462  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: PCT/US02/11950  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/286,782  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: 60/293,020  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/301,091  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/367,002  
; NUMBER OF SEQ ID NOS: 9  
; SEQ ID NO 1  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-945-853-1  
Query Match 100.0%; Score 96; DB 6; Length 188;  
Best Local Similarity 100.0%; Pred. No. 3.2e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 FRDSDIWPQEEPAIRPR 17  
Db 46 FRDSDIWPQEEPAIRPR 62  
|||||

RESULT 3  
US-10-945-853-2  
; Sequence 2, Application US/10945853  
; Publication No. US20050255117A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc.  
; APPLICANT: Sanicola-Nadel, Michele  
; APPLICANT: Adkins, Heather  
; APPLICANT: Miklasz, Steven Donald  
; APPLICANT: Rayhorn, Paul  
; APPLICANT: Schiffer, Susan Gail  
; APPLICANT: Williams, Kevin  
; TITLE OF INVENTION: Crypto-Specific Antibodies  
; FILE REFERENCE: BGN117CPCCN  
; CURRENT FILING DATE: 2004-09-20  
; PRIOR FILING DATE: 2004-09-20  
; PRIOR APPLICATION NUMBER: PCT/US02/31462  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: PCT/US02/11950  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/286,782  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: 60/293,020  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/301,091  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/367,002  
; NUMBER OF SEQ ID NOS: 9  
; SEQ ID NO 3  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-945-853-3  
Query Match 100.0%; Score 96; DB 6; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.2e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 FRDSDIWPQEEPAIRPR 17  
Db 1 FRDSDIWPQEEPAIRPR 17  
|||||

```
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 2
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-945-853-2

Query Match      100.0%; Score 96; DB 6; Length 188;
Best Local Similarity 100.0%; Pred. No. 3.2e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWPOEPAIRPR 17
Db 46 FRDSIWPOEPAIRPR 62

RESULT 4
US-11-000-463-252
; Sequence 252, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; PRIOR FILING DATE: 2004-11-29
; PRIOR FILING DATE: 2002-11-08
; PRIOR FILING DATE: 2002-11-08
; PRIOR FILING DATE: 2001-01-25
; PRIOR FILING DATE: 2001-01-25
; PRIOR FILING DATE: 2001-08-03
; PRIOR FILING DATE: 2000-01-25
; PRIOR FILING DATE: 2000-01-25
; PRIOR FILING DATE: 2000-07-17
; PRIOR FILING DATE: 2000-07-17
; PRIOR FILING DATE: 2000-08-03
; PRIOR FILING DATE: 2000-09-15
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 252
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-252

Query Match      45.8%; Score 44; DB 7; Length 115;
Best Local Similarity 57.1%; Pred. No. 2.3;
Matches 8; Conservative 1; Mismatches 5; Indels 5; Gaps 0;

Qy 3 DDSIWPOEPAIRPR 16
Db 4 DPASWPOEPAIRPR 17

RESULT 5
US-11-064-174-29
; Sequence 29, Application US/11064174
; Publication No. US20050282252A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
```

```
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELLS
; FILE REFERENCE: 09596-42U2
; CURRENT APPLICATION NUMBER: US/11/064,174
; PRIOR FILING DATE: 2005-02-22
; PRIOR FILING DATE: 1999-01-29
; PRIOR FILING DATE: 1998-04-10
; PRIOR FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain F01
US-11-064-174-29

Query Match      41.7%; Score 40; DB 7; Length 107;
Best Local Similarity 41.2%; Pred. No. 9.1;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 FRDSIWPOEPAIRPR 17
Db 28 FRDLGWYQKPGKAPK 44

RESULT 6
US-10-467-657-156
; Sequence 156, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: CONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 156
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-156

Query Match      41.1%; Score 39.5; DB 6; Length 228;
Best Local Similarity 45.0%; Pred. No. 22;
Matches 9; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

Qy 1 FRD-----DSIWPOEPAIR 15
Db 50 FKDRNRLYDSIQPTNPALK 69

RESULT 7
US-10-467-657-6148
; Sequence 6148, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
```



; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: Seqwin99, version 1.04  
; SEQ ID NO 6148  
; LENGTH: 255  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-6148

Query Match 41.1%; Score 39.5; DB 6; Length 255;  
Best Local Similarity 45.0%; Pred. No. 25;  
Matches 9; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

Qy 1 FRD-----DSTWPOEEPAIR 15  
Db 77 FKDRRLYDSIQPTNPALK 96

## RESULT 8

US-10-467-657-7562  
; Sequence 7562, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SpA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: Seqwin99, version 1.04  
; SEQ ID NO 7562  
; LENGTH: 255  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-7562

Query Match 41.1%; Score 39.5; DB 6; Length 255;  
Best Local Similarity 45.0%; Pred. No. 25;  
Matches 9; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

Qy 1 FRD-----DSTWPOEEPAIR 15  
Db 77 FKDRRLYDSIQPTNPALK 96

## RESULT 9

US-10-520-820-1  
; Sequence 1, Application US/10520820  
; Publication No. US2006000393A1  
; GENERAL INFORMATION:  
; APPLICANT: MUTABILIS S.A.  
; TITLE OF INVENTION: Pathogenicity determinants which can be used as targets for development of vaccines  
; TITLE OF INVENTION: means for preventing and controlling bacterial infections and/or  
; TITLE OF INVENTION: dissemination  
; FILE REFERENCE: 1621  
; CURRENT APPLICATION NUMBER: US/10/520,820  
; CURRENT FILING DATE: 2005-01-07  
; PRIOR APPLICATION NUMBER: PCT/EP2003/008209  
; PRIOR FILING DATE: 2003-07-09  
; PRIOR APPLICATION NUMBER: FR 0208636  
; PRIOR FILING DATE: 2002-07-09

; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 305  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-520-820-1

Query Match 40.6%; Score 39; DB 6; Length 305;  
Best Local Similarity 66.7%; Pred. No. 35;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RDDSINPQEE 10  
Db 171 RDDKHWPPEE 179

## RESULT 10

US-10-055-877-4  
; Sequence 4, Application US/10055877  
; Publication No. US20050288241A1  
; GENERAL INFORMATION:  
; APPLICANT: DeCristofaro, Marc  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Miller, Charles  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Zhong, Mei  
; APPLICANT: Anderson, David  
; APPLICANT: Ballinger, Robert  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Ratelli, Luca  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Andrew, David  
; APPLICANT: Mezes, Peter  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Eisen, Andrew  
; APPLICANT: Wolenc, Adam  
; APPLICANT: Baumgartner, Jason  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Vernet, Corine  
; APPLICANT: Taupier Jr., Raymond  
; APPLICANT: Pena, Carol  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Li, Li  
; APPLICANT: Casman, Stacie  
; APPLICANT: Boldog, Ference  
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby  
; FILE REFERENCE: 21402-251  
; CURRENT APPLICATION NUMBER: US/10/055,877  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: 60/262,892  
; PRIOR FILING DATE: 2001-01-19  
; PRIOR APPLICATION NUMBER: 60/263,598  
; PRIOR FILING DATE: 2001-01-23  
; PRIOR APPLICATION NUMBER: 60/263,799  
; PRIOR FILING DATE: 2001-01-24  
; PRIOR APPLICATION NUMBER: 60/264,117  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/264,139  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/264,478  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/263,351  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 60/272,870  
; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: 60/275,990  
; PRIOR FILING DATE: 2001-03-14



; PRIOR APPLICATION NUMBER: 60/275,927  
; PRIOR FILING DATE: 2001-03-14  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 512  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 441  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-055-877-4

Query Match 40.6%; Score 39; DB 6; Length 441;  
Best Local Similarity 75.0%; Pred. No. 51;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 WQDEPAI 14  
Db 106 WQDEPDI 113

RESULT 11

US-10-055-877-135  
; Sequence 135, Application US/10055877  
; Publication No. US20050288241A1  
; GENERAL INFORMATION:  
; APPLICANT: DeCristofaro, Marc  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Miller, Charles  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Zhong, Mei  
; APPLICANT: Anderson, David  
; APPLICANT: Ballinger, Robert  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Rattelli, Luca  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Andrew, David  
; APPLICANT: Mezes, Peter  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Eisen, Andrew  
; APPLICANT: Wolenc, Adam  
; APPLICANT: Baumgartner, Jason  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Vernet, Corine  
; APPLICANT: Taupier Jr., Raymond  
; APPLICANT: Pena, Carol  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Li, Li  
; APPLICANT: Casman, Stacie  
; APPLICANT: Boldog, Ference  
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby  
; FILE REFERENCE: 21402-251  
; CURRENT APPLICATION NUMBER: US/10/055,877  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: 60/262,892  
; PRIOR FILING DATE: 2001-01-19  
; PRIOR APPLICATION NUMBER: 60/263,598  
; PRIOR FILING DATE: 2001-01-23  
; PRIOR APPLICATION NUMBER: 60/264,117  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/264,139  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/264,478  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 60/272,870

; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: 60/275,990  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/275,927  
; PRIOR FILING DATE: 2001-03-14  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 512  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 135  
; LENGTH: 441  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-055-877-135

Query Match 40.6%; Score 39; DB 6; Length 441;  
Best Local Similarity 75.0%; Pred. No. 51;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 WQDEPAI 14  
Db 103 WQDEPDI 110

RESULT 12

US-10-055-877-134  
; Sequence 134, Application US/10055877  
; Publication No. US20050288241A1  
; GENERAL INFORMATION:  
; APPLICANT: DeCristofaro, Marc  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Miller, Charles  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Zhong, Mei  
; APPLICANT: Anderson, David  
; APPLICANT: Ballinger, Robert  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Rattelli, Luca  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Andrew, David  
; APPLICANT: Mezes, Peter  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Eisen, Andrew  
; APPLICANT: Wolenc, Adam  
; APPLICANT: Baumgartner, Jason  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Vernet, Corine  
; APPLICANT: Taupier Jr., Raymond  
; APPLICANT: Pena, Carol  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Li, Li  
; APPLICANT: Casman, Stacie  
; APPLICANT: Boldog, Ference  
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby  
; FILE REFERENCE: 21402-251  
; CURRENT APPLICATION NUMBER: US/10/055,877  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: 60/262,892  
; PRIOR FILING DATE: 2001-01-19  
; PRIOR APPLICATION NUMBER: 60/263,598  
; PRIOR FILING DATE: 2001-01-23  
; PRIOR APPLICATION NUMBER: 60/264,117  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/264,139  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/264,478  
; PRIOR FILING DATE: 2001-01-26

```
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 134
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-055-877-134

Query Match          40.6%; Score 39; DB 6; Length 479;
Best Local Similarity 75.0%; Pred. No. 55;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      7 WPOEPAI 14
Db      103 WPOEDPDI 110

RESULT 13
US-10-055-877-133
; Sequence 133, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: DeCristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shimkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
```

```
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 133
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-055-877-133

Query Match          40.6%; Score 39; DB 6; Length 485;
Best Local Similarity 75.0%; Pred. No. 55;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      7 WPOEPAI 14
Db      103 WPOEDPDI 110

RESULT 14
US-10-055-877-131
; Sequence 131, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: DeCristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eisen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shimkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
```

; PRIOR APPLICATION NUMBER: 60/264,117  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/264,139  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/264,478  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/263,351  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 60/272,870  
; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: 60/275,990  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/275,927  
; PRIOR FILING DATE: 2001-03-14  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 512  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 131  
; LENGTH: 637  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-055-877-131

Query Match 40.6%; Score 39; DB 6; Length 637;  
Best Local Similarity 75.0%; Pred. No. 72;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 7 WPOEPAI 14  
|||:|  
Db 254 WPQEDPDI 261

RESULT 15  
US-11-077-886-34  
; Sequence 34, Application US/11077886  
; Publication No. US20050266436A1  
; GENERAL INFORMATION:  
; APPLICANT: Sobek, Harald  
; APPLICANT: Frey, Bruno  
; APPLICANT: Antranikian, Garabed  
; APPLICANT: Boehlke, Kristina  
; APPLICANT: Pisanl, Francesca Maria  
; APPLICANT: Rossi, Mose  
; TITLE OF INVENTION: Mutant B-type DNA Polymerases Exhibiting Improved Performance in  
; FILE REFERENCE: 5328  
; CURRENT APPLICATION NUMBER: US/11/077,886  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: US/09/803,165  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: EP/00105155.6  
; PRIOR FILING DATE: 2000-03-11  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 34  
; LENGTH: 774  
; TYPE: PRT  
; ORGANISM: T. aggregans  
US-11-077-886-34

Query Match 40.6%; Score 39; DB 7; Length 774;  
Best Local Similarity 40.0%; Pred. No. 87;  
Matches 8; Conservative 3; Mismatches 3; Indels 6; Gaps 1;  
  
Qy 2 RDSIW-----POEPAIR 15  
|||:|  
Db 78 RDVEVVKLIFEHPQDVPALR 97

RESULT 16  
US-11-171-701-6  
; Sequence 6, Application US/11171701  
; Publication No. US2006008879A1  
; GENERAL INFORMATION:

; APPLICANT: Otani, Suzanne  
; APPLICANT: Blinkovsky, Alexander  
; APPLICANT: Ge, Haiyan  
; APPLICANT: Harris, Paul  
; APPLICANT: Yaver, Debbie  
; TITLE OF INVENTION: Polypeptides Having Alpha-Glucosidase Activity And  
; TITLE OF INVENTION: Polynucleotides Encoding Same  
; FILE REFERENCE: 10655.200-US  
; CURRENT APPLICATION NUMBER: US/11/171,701  
; CURRENT FILING DATE: 2005-06-29  
; PRIOR APPLICATION NUMBER: 60/585,336  
; PRIOR FILING DATE: 2005-06-29  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 6  
; LENGTH: 988  
; TYPE: PRT  
; ORGANISM: Aspergillus fumigatus  
US-11-171-701-6

Query Match 40.6%; Score 39; DB 7; Length 988;  
Best Local Similarity 50.0%; Pred. No. 11e+02;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
  
Qy 2 RDSIWPOEPAIRPR 17  
|||:|  
Db 853 RGGSLPMQEPALTTR 868

RESULT 17  
US-10-503-575-83  
; Sequence 83, Application US/10503575  
; Publication No. US20050244823A1  
; GENERAL INFORMATION:  
; APPLICANT: Drijfhout, Jan Wouter  
; APPLICANT: van Veeelen, Petrus Antonius  
; APPLICANT: Koning, Frits  
; TITLE OF INVENTION: NOVEL EPITOPES FOR CELIAC DISEASE AND AUTOIMMUNE DISEASES, METHOD:  
; FILE REFERENCE: 2799/72843-PCT-US  
; CURRENT APPLICATION NUMBER: US/10/503,575  
; CURRENT FILING DATE: 2004-08-04  
; PRIOR APPLICATION NUMBER: PCT/NL03/00077  
; PRIOR FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: EP 02075456.0  
; PRIOR FILING DATE: 2002-02-04  
; NUMBER OF SEQ ID NOS: 340  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 83  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-503-575-83

Query Match 39.6%; Score 38; DB 6; Length 17;  
Best Local Similarity 45.5%; Pred. No. 3.2;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 7 WPOEPAIRPR 17  
|||:|  
Db 4 WPOQPFPPQPQ 14

RESULT 18  
US-10-467-657-7960  
; Sequence 7960, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SpA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta

; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
 ; FILE REFERENCE:  
 ; CURRENT APPLICATION NUMBER: US/10/467,657  
 ; CURRENT FILING DATE: 2003-08-11  
 ; PRIOR APPLICATION NUMBER: GB-0103424.8  
 ; PRIOR FILING DATE: 2001-02-12  
 ; NUMBER OF SEQ ID NOS: 9218  
 ; SOFTWARE: SeqWin99, version 1.04  
 ; SEQ ID NO 7960  
 ; LENGTH: 92  
 ; TYPE: PRT  
 ; ORGANISM: Neisseria gonorrhoeae  
 ; US-10-467-657-7960

Query Match 39.6%; Score 38; DB 6; Length 92;  
 Best Local Similarity 50.0%; Pred. No. 16;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 6 IWQEEPAIR 15  
 Db 47 VWPKEADAVR 56

RESULT 19  
 US-11-092-353-2  
 ; Sequence 2, Application US/11092353  
 ; Publication No. US20050272653A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Knopf, John  
 ; APPLICANT: Seehra, Jasbir  
 ; TITLE OF INVENTION: BMP-3 PROPEPTIDES AND RELATED METHODS  
 ; FILE REFERENCE: BHPH-P01-002  
 ; CURRENT APPLICATION NUMBER: US/11/092,353  
 ; CURRENT FILING DATE: 2005-03-28  
 ; PRIOR APPLICATION NUMBER: US 60/557,100  
 ; PRIOR FILING DATE: 2004-03-26  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 335  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-11-092-353-2

Query Match 39.6%; Score 38; DB 7; Length 335;  
 Best Local Similarity 35.3%; Pred. No. 56;  
 Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 FRDDSIWQEEPAIRPR 17  
 Db 285 FKHQLWPSPPFRALKPR 301

RESULT 20  
 US-11-092-353-4  
 ; Sequence 4, Application US/11092353  
 ; Publication No. US20050272653A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Knopf, John  
 ; APPLICANT: Seehra, Jasbir  
 ; TITLE OF INVENTION: BMP-3 PROPEPTIDES AND RELATED METHODS  
 ; FILE REFERENCE: BHPH-P01-002  
 ; CURRENT APPLICATION NUMBER: US/11/092,353  
 ; CURRENT FILING DATE: 2005-03-28  
 ; PRIOR APPLICATION NUMBER: US 60/557,100  
 ; PRIOR FILING DATE: 2004-03-26  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 478  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-11-092-353-4

Query Match 39.6%; Score 38; DB 7; Length 478;  
 Best Local Similarity 35.3%; Pred. No. 78;  
 Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 FRDDSIWQEEPAIRPR 17  
 Db 318 FKHQLWPSPPFRALKPR 334

RESULT 21  
 US-10-453-372-192  
 ; Sequence 192, Application US/104533372  
 ; Publication No. US2006000323A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alsobrook, et al.  
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
 ; FILE REFERENCE: 21402-589 A  
 ; CURRENT APPLICATION NUMBER: US/10/453,372  
 ; CURRENT FILING DATE: 2003-06-03  
 ; PRIOR APPLICATION NUMBER: 09/789390  
 ; PRIOR FILING DATE: 2001-02-23  
 ; PRIOR APPLICATION NUMBER: 60/185967  
 ; PRIOR FILING DATE: 2000-03-01  
 ; PRIOR APPLICATION NUMBER: 09/823187  
 ; PRIOR FILING DATE: 2001-03-29  
 ; PRIOR APPLICATION NUMBER: 60/195792  
 ; PRIOR FILING DATE: 2000-03-10  
 ; PRIOR APPLICATION NUMBER: 09/839446  
 ; PRIOR FILING DATE: 2001-03-19  
 ; PRIOR APPLICATION NUMBER: 60/199476  
 ; PRIOR FILING DATE: 2000-03-25  
 ; PRIOR APPLICATION NUMBER: 09/863776  
 ; PRIOR FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: 60/208263  
 ; PRIOR FILING DATE: 2000-05-31  
 ; PRIOR APPLICATION NUMBER: 09/939398  
 ; PRIOR FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: 60/227800  
 ; PRIOR FILING DATE: 2000-08-25  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 1609  
 ; SOFTWARE: CuraSeqList version 0.1  
 ; SEQ ID NO 192  
 ; LENGTH: 2050  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (367)..(367)  
 ; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid  
 ; US-10-453-372-192

Query Match 39.6%; Score 38; DB 6; Length 2050;  
 Best Local Similarity 40.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 RDDSIIWQEEPAIRP 16  
 Db 889 QDDKNWDEDEPICP 903

RESULT 22  
 US-10-453-372-194  
 ; Sequence 194, Application US/104533372  
 ; Publication No. US2006000323A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alsobrook, et al.  
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
 ; FILE REFERENCE: 21402-589 A  
 ; CURRENT APPLICATION NUMBER: US/10/453,372  
 ; CURRENT FILING DATE: 2003-06-03  
 ; PRIOR APPLICATION NUMBER: 09/789390

```
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 194
; LENGTH: 3568
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-194
```

```
Query Match          39.6%; Score 38; DB 6; Length 3568;
Best Local Similarity 40.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      2 RDSIWPQEPPIR 16
      :||| :|||
Db      2813 QDDKNWDEPICIP 2827
```

```
RESULT 23
US-10-453-372-178
; Sequence 178, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Alsobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 178
; LENGTH: 3570
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-453-372-178
```

```
Query Match          39.6%; Score 38; DB 6; Length 3570;
Best Local Similarity 40.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      2 RDSIWPQEPPIR 16
      :||| :|||
Db      2815 QDDKNWDEPICIP 2829
```

```
RESULT 24
US-10-453-372-196
; Sequence 196, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Alsobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 196
; LENGTH: 3570
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-196
```

```
Query Match          39.6%; Score 38; DB 6; Length 3570;
Best Local Similarity 40.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      2 RDSIWPQEPPIR 16
      :||| :|||
Db      2815 QDDKNWDEPICIP 2829
```

```
RESULT 25
US-10-453-372-198
; Sequence 198, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Alsobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
```

; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 09/823187  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/195792  
; PRIOR FILING DATE: 2000-03-10  
; PRIOR APPLICATION NUMBER: 09/839446  
; PRIOR FILING DATE: 2001-03-19  
; PRIOR APPLICATION NUMBER: 60/199476  
; PRIOR FILING DATE: 2000-03-25  
; PRIOR APPLICATION NUMBER: 09/863776  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: 60/208263  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 09/939398  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: 60/227800  
; PRIOR FILING DATE: 2000-08-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1609  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 198  
; LENGTH: 3570  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-453-372-198

Query Match 39.6%; Score 38; DB 6; Length 3570;  
Best Local Similarity 40.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 RDDSIMPQEEPAIRP 16  
Db 2815 QDDKNWDEDEPICIP 2829

RESULT 26  
US-10-453-372-200  
; Sequence 200, Application US/10453372  
; Publication No. US2006000323A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsbrook, et al.  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-589 A  
; CURRENT APPLICATION NUMBER: US/10/453,372  
; PRIOR FILING DATE: 2003-06-03  
; PRIOR APPLICATION NUMBER: 09/789390  
; PRIOR FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: 60/185967  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 09/823187  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/195792  
; PRIOR FILING DATE: 2000-03-10  
; PRIOR APPLICATION NUMBER: 09/839446  
; PRIOR FILING DATE: 2001-03-19  
; PRIOR APPLICATION NUMBER: 60/199476  
; PRIOR FILING DATE: 2000-03-25  
; PRIOR APPLICATION NUMBER: 09/863776  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: 60/208263  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 09/939398  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: 60/227800  
; PRIOR FILING DATE: 2000-08-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1609  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 200  
; LENGTH: 3570  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-453-372-200

Query Match 39.6%; Score 38; DB 6; Length 3570;  
Best Local Similarity 40.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 RDDSIMPQEEPAIRP 16  
Db 2815 QDDKNWDEDEPICIP 2829

RESULT 27  
US-10-453-372-202  
; Sequence 202, Application US/10453372  
; Publication No. US2006000323A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsbrook, et al.  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-589 A  
; CURRENT APPLICATION NUMBER: US/10/453,372  
; PRIOR FILING DATE: 2003-06-03  
; PRIOR APPLICATION NUMBER: 09/789390  
; PRIOR FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: 60/185967  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 09/823187  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/195792  
; PRIOR FILING DATE: 2000-03-10  
; PRIOR APPLICATION NUMBER: 09/839446  
; PRIOR FILING DATE: 2001-03-19  
; PRIOR APPLICATION NUMBER: 60/199476  
; PRIOR FILING DATE: 2000-03-25  
; PRIOR APPLICATION NUMBER: 09/863776  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: 60/208263  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 09/939398  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: 60/227800  
; PRIOR FILING DATE: 2000-08-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1609  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 202  
; LENGTH: 3570  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-453-372-202

Query Match 39.6%; Score 38; DB 6; Length 3570;  
Best Local Similarity 40.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 RDDSIMPQEEPAIRP 16  
Db 2815 QDDKNWDEDEPICIP 2829

RESULT 28  
US-10-453-372-204  
; Sequence 204, Application US/10453372  
; Publication No. US2006000323A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsbrook, et al.  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-589 A  
; CURRENT APPLICATION NUMBER: US/10/453,372  
; PRIOR FILING DATE: 2003-06-03  
; PRIOR APPLICATION NUMBER: 09/789390  
; PRIOR FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: 60/185967  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 09/823187

```
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 204
; LENGTH: 3570
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-204

Query Match      39.6%; Score 38; DB 6; Length 3570;
Best Local Similarity 40.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      2 RDSIWQPQEPAIRP 16
       :||| :||| :|||
Db      2815 QDDKNWDEDEPICIP 2829

RESULT 29
US-10-453-372-206
; Sequence 206, Application US/10453372
; Publication No. US200600323A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 206
; LENGTH: 3570
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-206

Query Match      39.6%; Score 38; DB 6; Length 3570;
Best Local Similarity 40.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      2 RDSIWQPQEPAIRP 16
       :||| :||| :|||
Db      2815 QDDKNWDEDEPICIP 2829

Best Local Similarity 40.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      2 RDSIWQPQEPAIRP 16
       :||| :||| :|||
Db      2815 QDDKNWDEDEPICIP 2829

RESULT 30
US-11-155-775-56
; Sequence 56, Application US/11155775
; Publication No. US20050266013A1
; GENERAL INFORMATION:
; APPLICANT: Morell, Andreas
; APPLICANT: Imboden, Martin
; APPLICANT: Stadler, Beda
; APPLICANT: Miescher, Sylvia
; APPLICANT: Vogel, Monique
; APPLICANT: Amstutz, Hanspeter
; TITLE OF INVENTION: POLYPEPTIDES CAPABLE OF FORMING ANTIGEN BINDING
; TITLE OF INVENTION: STRUCTURES WITH SPECIFICITY FOR THE RHESUS D ANTIGENS,
; TITLE OF INVENTION: THE DNA ENCODING THEM AND THE PROCESS FOR THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; FILE REFERENCE: 6816/P63221US0
; CURRENT APPLICATION NUMBER: US/11/155,775
; CURRENT FILING DATE: 2005-06-20
; PRIOR APPLICATION NUMBER: US/09/147,443
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: PCT/EP97/03253
; PRIOR FILING DATE: 1997-06-20
; PRIOR APPLICATION NUMBER: EP 96810421.6
; PRIOR FILING DATE: 1996-06-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-775-56

Query Match      38.5%; Score 37; DB 7; Length 105;
Best Local Similarity 37.5%; Pred. No. 26;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY      2 RDSIWQPQEPAIRP 17
       :||| :||| :|||
Db      28 RNDLTWYQQKPKAPK 43

RESULT 31
US-10-925-366A-153
; Sequence 153, Application US/10925366A
; Publication No. US20050271663A1
; GENERAL INFORMATION:
; APPLICANT: Ignatovich, Olga
; APPLICANT: Dewilddt, Rudolph M.T.
; APPLICANT: Benjamin, Woolven
; APPLICANT: Grant, Steven
; APPLICANT: Jones, Philip
; APPLICANT: Basran, Amrik
; APPLICANT: Brewis, Neil
; TITLE OF INVENTION: Compositions and Methods for Treating Inflammatory Disorders
; FILE REFERENCE: 8039/2105
; CURRENT APPLICATION NUMBER: US/10/925,366A
; CURRENT FILING DATE: 2004-08-24
; PRIOR APPLICATION NUMBER: US 10/744,774
; PRIOR FILING DATE: 2003-12-23
; PRIOR APPLICATION NUMBER: PCT/GB2003/002804
; PRIOR FILING DATE: 2003-06-30
; PRIOR APPLICATION NUMBER: PCT/GB2002/03014
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: GB 0230202.4
; PRIOR FILING DATE: 2002-12-27
```

RESULT 35  
US-10-454-437-386  
; Sequence 386, Application US/10454437  
; Publication No. US2005027115A1  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Markus  
; APPLICANT: Kroger, Burkhard



```
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberkauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128GFCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 386
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-386

Query Match      38.5%; Score 37; DB 6; Length 373;
Best Local Similarity 50.0%; Pred. No. 88;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy      5 S1NPQEEP 12
Db      232 AVMPEKP 239

RESULT 36
US-11-024-959-301
; Sequence 301, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 301
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
US-11-024-959-301

Query Match      38.0%; Score 36.5; DB 7; Length 502;
```

```
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy      3 DDS1NPQEEP-AIR 15
Db      25 DVELWPKEAPKAVR 38

RESULT 37
US-11-054-669-83
; Sequence 83, Application US/11054669
; Publication No. US20050261480A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: SUPER HUMANIZED ANTIBODIES
; FILE REFERENCE: 30219/US/3
; CURRENT APPLICATION NUMBER: US/11/054,669
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US 10/194,975
; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 83
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-054-669-83

Query Match      37.5%; Score 36; DB 7; Length 96;
Best Local Similarity 45.5%; Pred. No. 34;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      7 WPOEPAIRPR 17
Db      36 WYQKFGGLAPR 46

RESULT 38
US-11-084-554-105
; Sequence 105, Application US/11084554
; Publication No. US20050260679A1
; GENERAL INFORMATION:
; APPLICANT: Kellermann, Sirid-Ai
; APPLICANT: Green, Larry L.
; APPLICANT: Korver, Wouter
; TITLE OF INVENTION: REDUCING THE RISK OF HUMAN ANTI-HUMAN
; FILE REFERENCE: ABGENIX.100A
; CURRENT APPLICATION NUMBER: US/11/084,554
; CURRENT FILING DATE: 2005-03-17
; PRIOR APPLICATION NUMBER: 60/554,372
; PRIOR FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: 60/574,661
; PRIOR FILING DATE: 2004-05-24
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-084-554-105

Query Match      37.5%; Score 36; DB 7; Length 96;
Best Local Similarity 45.5%; Pred. No. 34;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      7 WPOEPAIRPR 17
Db      36 WYQKFGGLAPR 46
```

## RESULT 39

US-11-004-590-93  
; Sequence 93, Application US/11004590  
; Publication No. US20060008883A1  
; GENERAL INFORMATION:  
; APPLICANT: Lazar, Gregory Alan  
; APPLICANT: Desjarlais, John W.  
; APPLICANT: Hammond, Phillip W.  
; TITLE OF INVENTION: METHODS OF GENERATING VARIANT PROTEINS WITH INCREASED HOST STRING  
; TITLE OF INVENTION: CONTENT AND COMPOSITIONS THEREOF  
; FILE REFERENCE: 185832/US/5  
; CURRENT APPLICATION NUMBER: US/11/004,590  
; CURRENT FILING DATE: 2004-12-03  
; PRIOR APPLICATION NUMBER: US 60/527,167  
; PRIOR FILING DATE: 2003-12-04  
; PRIOR APPLICATION NUMBER: US 60/581,613  
; PRIOR FILING DATE: 2004-06-21  
; PRIOR APPLICATION NUMBER: US 60/601,665  
; PRIOR FILING DATE: 2004-08-13  
; PRIOR APPLICATION NUMBER: US 60/619,483  
; PRIOR FILING DATE: 2004-10-14  
; NUMBER OF SEQ ID NOS: 458  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 93  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-004-590-93

Query Match 37.5%; Score 36; DB 7; Length 96;  
Best Local Similarity 45.5%; Pred. No. 34;  
Matches 5; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

Qy 7 WQEQPAIRPR 17  
| | : | : |  
Db 36 WYQKPGGLAPR 46

## RESULT 40

US-11-136-250-105  
; Sequence 105, Application US/11136250  
; Publication No. US20060021074A1  
; GENERAL INFORMATION:  
; APPLICANT: Kellermann, Sirid-Ai  
; APPLICANT: Green, Larry L.  
; APPLICANT: Korver, Wouter  
; TITLE OF INVENTION: REDUCING THE RISK OF HUMAN ANTI-HUMAN  
; TITLE OF INVENTION: ANTIBODIES THROUGH V GENE MANIPULATION  
; FILE REFERENCE: AGENIX.100A2  
; CURRENT APPLICATION NUMBER: US/11/136,250  
; CURRENT FILING DATE: 2005-05-23  
; PRIOR APPLICATION NUMBER: 11/084,554  
; PRIOR FILING DATE: 2005-03-17  
; PRIOR APPLICATION NUMBER: PCT/US2005/009306  
; PRIOR FILING DATE: 2005-03-17  
; PRIOR APPLICATION NUMBER: 60/574,661  
; PRIOR FILING DATE: 2004-05-24  
; PRIOR APPLICATION NUMBER: 60/554,372  
; PRIOR FILING DATE: 2004-03-19  
; NUMBER OF SEQ ID NOS: 266  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 105  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-136-250-105

Query Match 37.5%; Score 36; DB 7; Length 96;  
Best Local Similarity 45.5%; Pred. No. 34;  
Matches 5; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

Qy 7 WQEQPAIRPR 17  
| | : | : |

## Db

36 WYQKPGGLAPR 46

## RESULT 41

US-10-771-257-69  
; Sequence 69, Application US/10771257  
; Publication No. US2005028864A1  
; GENERAL INFORMATION:  
; APPLICANT: Medical Research Council  
; APPLICANT: SISSA - Scuola Superiore Internazionale di Studi Avanzati  
; APPLICANT: Cattaneo, Antonino  
; APPLICANT: Maritan, Amos  
; APPLICANT: Visintin, Michela  
; APPLICANT: Rabbitts, Terrence H  
; APPLICANT: Settanni, Giovanni  
; TITLE OF INVENTION: Intracellular antibodies  
; FILE REFERENCE: 18396/2272  
; CURRENT APPLICATION NUMBER: US/10/771,257  
; CURRENT FILING DATE: 2004-02-03  
; PRIOR APPLICATION NUMBER: PCT/GB02/03512  
; PRIOR FILING DATE: 2002-08-01  
; PRIOR APPLICATION NUMBER: GB 0119004.0  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: GB 0121577.1  
; PRIOR FILING DATE: 2001-09-06  
; PRIOR APPLICATION NUMBER: GB 0200928.0  
; PRIOR FILING DATE: 2002-01-16  
; PRIOR APPLICATION NUMBER: GB 0203569.9  
; PRIOR FILING DATE: 2002-02-14  
; PRIOR APPLICATION NUMBER: IT RM2001A000633  
; PRIOR FILING DATE: 2001-10-25  
; NUMBER OF SEQ ID NOS: 124  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 69  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-771-257-69

Query Match 37.5%; Score 36; DB 6; Length 108;  
Best Local Similarity 33.3%; Pred. No. 38;  
Matches 5; Conservative 5; Mismatches 5; Indels 5; Gaps 0;

Qy 3 DDSIWQSEPAIRPR 17  
: | : | : | : |  
Db 31 NDLVWCQKPGKAPK 45

## RESULT 42

US-11-127-677-67  
; Sequence 67, Application US/11127677  
; Publication No. US20050272107A1  
; GENERAL INFORMATION:  
; APPLICANT: Medical Research Council  
; APPLICANT: Rabbitts, Terence H  
; APPLICANT: Tanaka, Tomoyuki  
; TITLE OF INVENTION: Intracellular antibodies  
; FILE REFERENCE: 18396/2462  
; CURRENT APPLICATION NUMBER: US/11/127,677  
; CURRENT FILING DATE: 2005-05-12  
; PRIOR APPLICATION NUMBER: PCT/GB03/04942  
; PRIOR FILING DATE: 2003-11-14  
; PRIOR APPLICATION NUMBER: GB 0226729.2  
; PRIOR FILING DATE: 2002-11-15  
; NUMBER OF SEQ ID NOS: 150  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 67  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Derived protein sequence of scFv  
US-11-127-677-67

Query Match 37.5%; Score 36; DB 7; Length 108;  
Best Local Similarity 33.3%; Pred. No. 38;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 DDSIWPOEPAIRPR 17  
DB 31 NDLVWCQKPKGKPK 45

## RESULT 43

US-10-821-234-869  
; Sequence 869, Application US/10821234  
; Publication No. US20050255114A1  
; GENERAL INFORMATION:  
; APPLICANT: Labat, Ivan  
; APPLICANT: Stache-Crain, Birgit  
; APPLICANT: Andarmani, Susan  
; APPLICANT: Tang, Y. Tom  
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
; FILE REFERENCE: 821A  
; CURRENT APPLICATION NUMBER: US/10/821,234  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: US 60/462,047  
; PRIOR FILING DATE: 2003-04-07  
; NUMBER OF SEQ ID NOS: 1704  
; SOFTWARE: pt\_seq\_genes Version 1.0  
; SEQ ID NO 869  
; LENGTH: 148  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-821-234-869

Query Match 37.5%; Score 36; DB 6; Length 148;  
Best Local Similarity 55.6%; Pred. No. 52;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 PQEPAIRP 16  
DB 128 PDRQPSLRP 136

## RESULT 44

US-11-054-515-1505  
; Sequence 1505, Application US/11054515  
; Publication No. US2005025532A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS  
; FILE REFERENCE: PF523P3  
; CURRENT APPLICATION NUMBER: US/11/054,515  
; CURRENT FILING DATE: 2005-02-10  
; PRIOR APPLICATION NUMBER: 60/543,296  
; PRIOR FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/580,347  
; PRIOR FILING DATE: 2004-06-18  
; PRIOR APPLICATION NUMBER: 10/293,418  
; PRIOR FILING DATE: 2002-11-14  
; PRIOR APPLICATION NUMBER: 60/331,469  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: 60/340,817  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 09/880,748  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 3247  
; SEQ ID NO 1505  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-054-515-1505

Query Match 37.5%; Score 36; DB 7; Length 252;  
Best Local Similarity 46.7%; Pred. No. 86;  
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 DDSIWPOEPAIRPR 17  
DB 173 DDVHWYQQLPGTAPR 187

## RESULT 45

US-11-156-084-195  
; Sequence 195, Application US/11156084  
; Publication No. US20060010515A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology LLC  
; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to  
; TITLE OF INVENTION: agronomically interesting phenotypes  
; FILE REFERENCE: (38-21)  
; CURRENT APPLICATION NUMBER: US/11/156,084  
; CURRENT FILING DATE: 2005-06-17  
; NUMBER OF SEQ ID NOS: 364  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 195  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Mycobacterium leprae  
US-11-156-084-195

Query Match 37.5%; Score 36; DB 7; Length 311;  
Best Local Similarity 46.7%; Pred. No. 1.1e+02;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 DDSIWPOEPAIRPR 17  
DB 112 DDWSFPGTDPVAVR 126

## RESULT 46

US-10-467-657-8554  
; Sequence 8554, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO 8554  
; LENGTH: 322  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-8554

Query Match 37.5%; Score 36; DB 6; Length 322;  
Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 RDSIWPOE 10

Db 187 RDSKLWPE 195  
|| :|||

RESULT 47  
US-11-152-697-4  
; Sequence 4, Application US/11152697  
; Publication No. US20060003367A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL HUMAN KUPFFER CELL RECEPTOR  
; FILE REFERENCE: D0242 NP  
; CURRENT APPLICATION NUMBER: US/11/152,697  
; PRIOR FILING DATE: 2005-06-14  
; PRIOR APPLICATION NUMBER: 60/580,006  
; PRIOR FILING DATE: 2004-06-15  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 328  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-152-697-4

Query Match 37.5%; Score 36; DB 7; Length 328;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 S1WPQEEP 12  
:|:|:|  
Db 20 SLWPREPP 27

RESULT 48  
US-10-641-678-58  
; Sequence 58, Application US/10641678  
; Publication No. US2005027172A1  
; GENERAL INFORMATION:  
; APPLICANT: Day, Anthony, G.  
; APPLICANT: Goedegebuur, Frits  
; APPLICANT: Gualfetti, Peter  
; APPLICANT: Mitchinson, Colin  
; APPLICANT: Neefe, Paulien  
; APPLICANT: Sandgren, Mats  
; APPLICANT: Shaw, Andrew  
; APPLICANT: Stahlberg, Jerry  
; TITLE OF INVENTION: Novel Variant Hypocrea jecorina CBH1  
; FILE REFERENCE: GC772-3  
; CURRENT APPLICATION NUMBER: US/10/641,678  
; CURRENT FILING DATE: 2003-08-15  
; PRIOR APPLICATION NUMBER: US 60/458,853  
; PRIOR FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: US 60/458,696  
; PRIOR FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: US 60/456,368  
; PRIOR FILING DATE: 2003-03-21  
; PRIOR APPLICATION NUMBER: US 60/404,063  
; PRIOR FILING DATE: 2002-08-16  
; NUMBER OF SEQ ID NOS: 77  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 58  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Leptosphaeria maculans  
US-10-641-678-58

Query Match 37.5%; Score 36; DB 6; Length 447;  
Best Local Similarity 46.2%; Pred. No. 1.5e+02;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 DSIWPQEP 16

Db 391 DSNWPREADPAKP 403  
|| |||:

RESULT 49  
US-11-087-100-20  
; Sequence 20, Application US/11087100  
; Publication No. US20050266440A1  
; GENERAL INFORMATION:  
; APPLICANT: Metz, James  
; APPLICANT: Barclay, William  
; APPLICANT: Platt, James  
; APPLICANT: Kuner, Jerry  
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthase  
; FILE REFERENCE: 2997-29  
; CURRENT APPLICATION NUMBER: US/11/087,100  
; CURRENT FILING DATE: 2005-03-21  
; PRIOR APPLICATION NUMBER: 09/231,899  
; PRIOR FILING DATE: 1999-01-14  
; PRIOR APPLICATION NUMBER: 60/284,066  
; PRIOR FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 60/298,796  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/323,269  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 20  
; LENGTH: 450  
; TYPE: PRT  
; ORGANISM: Schizochytrium sp.  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (370)..(370)  
; OTHER INFORMATION: The 'Xaa' at location 370 stands for Leu.  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (371)..(371)  
; OTHER INFORMATION: The 'Xaa' at location 371 stands for Ala, or Val.  
US-11-087-100-20

Query Match 37.5%; Score 36; DB 7; Length 450;  
Best Local Similarity 37.5%; Pred. No. 1.5e+02;  
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 FRDSIWPQEP 16  
:|:|:|:|:|  
Db 149 FKDASHWSEREQSNKP 164

RESULT 50  
US-11-087-100-28  
; Sequence 28, Application US/11087100  
; Publication No. US20050266440A1  
; GENERAL INFORMATION:  
; APPLICANT: Metz, James  
; APPLICANT: Barclay, William  
; APPLICANT: Platt, James  
; APPLICANT: Kuner, Jerry  
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthase  
; FILE REFERENCE: 2997-29  
; CURRENT APPLICATION NUMBER: US/11/087,100  
; CURRENT FILING DATE: 2005-03-21  
; PRIOR APPLICATION NUMBER: 09/231,899  
; PRIOR FILING DATE: 1999-01-14  
; PRIOR APPLICATION NUMBER: 60/284,066  
; PRIOR FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 60/298,796  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/323,269  
; PRIOR FILING DATE: 2001-09-18

```

; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 28
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Schizochytrium sp.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (370)..(370)
; OTHER INFORMATION: The 'Xaa' at location 370 stands for Leu.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (371)..(371)
; OTHER INFORMATION: The 'Xaa' at location 371 stands for Ala, or Val.
US-11-087-100-28

      Query Match          37.5%;      Score 36;      DB 7;      Length 450;
      Best Local Similarity 37.5%;      Pred. No. 1.5e+02;
      Matches- 6;      Conservative 4;      Mismatches 6;      Indels 0;      Gaps 0;

Oy      1      FRDSTIWQEEPAIRP 16
      |:|:|:|:|:|
Db      149      FKDAHWSEREQSNKP 164

```

Search completed: February 3, 2006, 19:02:23  
Job time : 14 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
 Copyright (c) 1993 - 2006 Bioceleration Ltd.  
 OM protein - protein search, using sw model  
 Run on: February 3, 2006, 18:57:01 ; Search time 119.5 Seconds  
 (without alignments)  
 59.440 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_46\_62  
 Perfect score: 96  
 Sequence: 1 FRDSDIWPQEPAIRPR 17

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Published Applications\_AA\_Main:  
 1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:  
 2: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:  
 3: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:  
 4: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep:  
 5: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep:  
 6: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	129	4	US-10-390-566-4
2	96	100.0	129	4	US-10-390-566-11
3	96	100.0	129	4	US-10-390-566-17
4	96	100.0	129	4	US-10-390-566-24
5	96	100.0	129	4	US-10-390-566-3
6	96	100.0	129	4	US-10-390-566-10
7	96	100.0	129	4	US-10-390-566-16
8	96	100.0	129	4	US-10-390-566-23
9	96	100.0	174	4	US-10-665-602-7
10	96	100.0	187	4	US-10-225-486-66
11	96	100.0	188	4	US-10-241-220-58
12	96	100.0	188	4	US-10-388-578-6
13	96	100.0	188	4	US-10-390-566-1
14	96	100.0	188	4	US-10-390-566-2
15	96	100.0	188	4	US-10-390-566-9
16	96	100.0	188	4	US-10-390-566-22
17	96	100.0	188	4	US-10-362-597A-3
18	96	100.0	188	4	US-10-362-597A-4
19	96	100.0	188	4	US-10-362-597A-6
20	96	100.0	188	4	US-10-362-597A-96
21	96	100.0	188	4	US-10-407-481-3
22	96	100.0	188	4	US-10-407-481-4
23	96	100.0	188	4	US-10-407-481-6
24	96	100.0	188	4	US-10-407-481-96
25	96	100.0	188	4	US-10-712-124-58
26	96	100.0	188	4	US-10-693-538-1
27	96	100.0	188	4	US-10-693-538-2

28	96	100.0	188	4	US-10-389-431-6	Sequence 6, Appli
29	96	100.0	188	4	US-10-816-476-3	Sequence 3, Appli
30	96	100.0	188	4	US-10-816-476-4	Sequence 4, Appli
31	96	100.0	188	4	US-10-816-476-6	Sequence 6, Appli
32	96	100.0	188	4	US-10-816-476-96	Sequence 96, Appli
33	96	100.0	188	5	US-10-872-972-58	Sequence 58, Appli
34	96	100.0	188	5	US-10-872-991-58	Sequence 58, Appli
35	96	100.0	188	5	US-10-491-997-28	Sequence 28, Appli
36	96	100.0	188	5	US-10-491-997-72	Sequence 72, Appli
37	96	100.0	188	5	US-10-940-431-2	Sequence 2, Appli
38	96	100.0	188	5	US-10-983-340-13	Sequence 13, Appli
39	96	100.0	360	4	US-10-390-566-7	Sequence 7, Appli
40	96	100.0	360	4	US-10-390-566-14	Sequence 14, Appli
41	96	100.0	360	4	US-10-390-566-20	Sequence 20, Appli
42	96	100.0	360	4	US-10-390-566-27	Sequence 27, Appli
43	96	100.0	367	4	US-10-390-566-6	Sequence 6, Appli
44	96	100.0	367	4	US-10-390-566-13	Sequence 13, Appli
45	96	100.0	367	4	US-10-390-566-19	Sequence 19, Appli
46	96	100.0	367	4	US-10-390-566-26	Sequence 26, Appli
47	64	66.7	160	4	US-10-016-447-7	Sequence 7, Appli
48	64	66.7	166	4	US-10-424-599-231675	Sequence 231675,
49	64	66.7	171	5	US-10-940-431-1	Sequence 1, Appli
50	54	56.2	9	4	US-10-362-597A-80	Sequence 80, Appli
51	54	56.2	9	4	US-10-407-481-80	Sequence 80, Appli
52	54	56.2	9	4	US-10-816-476-80	Sequence 80, Appli
53	48	50.0	123	4	US-10-424-599-197561	Sequence 197561,
54	48	50.0	266	4	US-10-425-115-282871	Sequence 282871,
55	48	50.0	496	4	US-10-369-493-6305	Sequence 6305, Ap
56	48	50.0	606	6	US-11-097-143-225	Sequence 225, App
57	48	50.0	606	6	US-11-097-143-228	Sequence 228, App
58	48	50.0	1081	4	US-10-451-467A-714	Sequence 714, App
59	47	49.0	18	4	US-10-362-597A-12	Sequence 12, Appl
60	47	49.0	18	4	US-10-407-481-12	Sequence 12, Appl
61	47	49.0	18	4	US-10-816-476-12	Sequence 12, Appl
62	47	49.0	96	4	US-10-767-701-36630	Sequence 36630, A
63	47	49.0	137	4	US-10-029-386-28993	Sequence 28993, A
64	47	49.0	216	4	US-10-425-114-55654	Sequence 55654, A
65	47	49.0	348	5	US-10-739-930-7342	Sequence 7342, Ap
66	47	49.0	349	4	US-10-425-115-319948	Sequence 319948,
67	46	47.9	293	4	US-10-425-115-202069	Sequence 202069,
68	46	47.9	988	5	US-10-647-268-2	Sequence 2, Appli
69	46	47.9	988	5	US-10-647-268-4	Sequence 4, Appli
70	46	47.9	988	5	US-10-647-268-10	Sequence 10, Appli
71	46	47.9	1003	5	US-10-647-268-6	Sequence 6, Appli
72	45	46.9	89	5	US-10-732-923-19883	Sequence 19883, A
73	45	46.9	90	5	US-10-732-923-19880	Sequence 19880, A
74	45	46.9	90	5	US-10-732-923-19881	Sequence 19881, A
75	45	46.9	91	5	US-10-732-923-19879	Sequence 19879, A
76	45	46.9	180	4	US-10-437-963-199813	Sequence 199813,
77	45	46.9	184	5	US-10-499-353A-520	Sequence 520, App
78	45	46.9	490	4	US-10-080-334-250	Sequence 250, App
79	45	46.9	490	4	US-10-108-260A-4335	Sequence 4335, Ap
80	45	46.9	490	4	US-10-275-595A-23	Sequence 23, Appli
81	44.5	46.4	562	5	US-10-450-763-45932	Sequence 45932, A
82	45	46.9	945	4	US-10-360-522-59	Sequence 59, Appli
83	45	46.9	1572	6	US-11-097-143-15396	Sequence 15396, A
84	44.5	46.4	906	4	US-10-369-493-6526	Sequence 6526, Ap
85	44	45.8	55	3	US-09-764-877-1474	Sequence 1474, Ap
86	44	45.8	55	4	US-10-242-515-1474	Sequence 1474, Ap
87	44	45.8	115	4	US-10-291-265-252	Sequence 252, App
88	44	45.8	608	4	US-10-282-122A-48913	Sequence 48913, A
89	44	45.8	988	5	US-10-755-966-12	Sequence 12, Appl
90	44	45.8	988	5	US-10-755-966-22	Sequence 22, Appli
91	44	45.8	1040	4	US-10-360-522-57	Sequence 57, Appli
92	44	45.8	1312	4	US-10-369-493-1950	Sequence 1950, Ap
93	43	44.8	99	3	US-09-864-408A-6900	Sequence 6900, Ap
94	43	44.8	167	4	US-10-106-698-4754	Sequence 4754, Ap
95	43	44.8	195	4	US-10-425-114-64958	Sequence 64958, A
96	43	44.8	222	4	US-10-425-115-223066	Sequence 223066,
97	43	44.8	335	4	US-10-282-122A-51097	Sequence 51097, A
98	43	44.8	356	4	US-10-156-761-13384	Sequence 13384, A
99	43	44.8	416	4	US-10-029-386-33334	Sequence 33334, A
100	43	44.8	620	4	US-10-108-260A-3458	Sequence 3458, Ap

101	43	44.8	663	4	US-10-108-260A-3492	Sequence 3492, Ap	174	42	43.8	970	5	US-10-755-966-8	Sequence 8, Appli
102	43	44.8	707	6	US-11-097-143-22388	Sequence 32388, A	175	42	43.8	979	4	US-10-360-522-55	Sequence 55, Appl
103	43	44.8	776	4	US-10-280-662-46	Sequence 46, Appl	176	42	43.8	979	4	US-10-360-522-58	Sequence 58, Appl
104	43	44.8	776	4	US-10-298-680-46	Sequence 46, Appl	177	42	43.8	979	5	US-10-755-966-2	Sequence 2, Appli
105	43	44.8	776	4	US-10-702-400-56	Sequence 56, Appl	178	42	43.8	1096	5	US-10-450-763-39265	Sequence 32665, A
106	43	44.8	776	5	US-10-805-650-56	Sequence 56, Appl	179	42	43.8	1427	4	US-10-282-122A-49525	Sequence 49525, A
107	43	44.8	1084	4	US-10-415-011-17	Sequence 17, Appl	180	42	43.8	1736	3	US-09-919-497-98	Sequence 98, Appl
108	43	44.8	1113	3	US-09-894-998-51	Sequence 51, Appl	181	42	43.8	1906	6	US-11-097-143-15201	Sequence 15201, A
109	43	44.8	1113	4	US-10-121-988-51	Sequence 51, Appl	182	41.5	43.2	82	4	US-10-425-115-365368	Sequence 365368, A
110	43	44.8	1113	4	US-10-200-562-51	Sequence 51, Appl	183	41.5	43.2	145	4	US-10-425-115-296515	Sequence 147391, A
111	43	44.8	1113	4	US-10-237-551-51	Sequence 51, Appl	184	41.5	43.2	747	4	US-10-437-963-147391	Sequence 728, App
112	43	44.8	1113	5	US-10-945-050-51	Sequence 51, Appl	185	41.5	43.2	944	4	US-10-451-167A-728	Sequence 259888, A
113	43	44.8	1114	4	US-10-200-562-202	Sequence 202, App	186	41	42.7	73	4	US-10-425-115-259888	Sequence 4408, Ap
114	43	44.8	1114	4	US-10-237-551-202	Sequence 202, App	187	41	42.7	104	4	US-10-108-260A-4408	Sequence 196591, A
115	43	44.8	1114	5	US-10-945-050-202	Sequence 202, App	188	41	42.7	124	4	US-10-424-599-196591	Sequence 330149, A
116	43	44.8	1231	4	US-10-052-648A-37	Sequence 37, Appl	189	41	42.7	124	4	US-10-425-115-330149	Sequence 330149, A
117	43	44.8	1234	4	US-10-052-648A-36	Sequence 36, Appl	190	41	42.7	169	4	US-10-276-774-1813	Sequence 1813, Ap
118	43	44.8	1243	4	US-10-196-935A-4	Sequence 4, Appli	191	41	42.7	178	4	US-10-437-963-183715	Sequence 183715, A
119	43	44.8	1243	4	US-10-478-146-8	Sequence 8, Appli	192	41	42.7	197	4	US-10-282-122A-53487	Sequence 53487, A
120	42.5	44.3	921	6	US-11-097-143-18348	Sequence 18348, A	193	41	42.7	227	5	US-10-450-763-33579	Sequence 33579, A
121	42	43.8	75	4	US-10-425-115-335674	Sequence 335674, A	194	41	42.7	227	5	US-10-450-763-52702	Sequence 52702, A
122	42	43.8	83	4	US-10-425-115-208701	Sequence 208701, A	195	41	42.7	227	5	US-11-097-143-10050	Sequence 10050, A
123	42	43.8	84	5	US-10-732-923-19885	Sequence 19885, A	196	41	42.7	301	6	US-11-097-143-38754	Sequence 38754, A
124	42	43.8	84	5	US-10-732-923-19890	Sequence 19890, A	197	41	42.7	448	5	US-10-450-763-56722	Sequence 56722, A
125	42	43.8	85	5	US-10-732-923-19929	Sequence 19929, A	198	41	42.7	474	4	US-10-282-122A-61516	Sequence 61516, A
126	42	43.8	86	5	US-10-732-923-19889	Sequence 19889, A	199	41	42.7	493	5	US-10-732-923-9693	Sequence 9693, Ap
127	42	43.8	87	5	US-10-732-923-19944	Sequence 19944, A	200	41	42.7	522	4	US-10-437-963-131694	Sequence 131694, A
128	42	43.8	89	5	US-10-732-923-19882	Sequence 19882, A	201	41	42.7	540	5	US-10-756-149-5586	Sequence 5586, Ap
129	42	43.8	90	5	US-10-732-923-19945	Sequence 19945, A	202	41	42.7	541	3	US-09-972-211-4	Sequence 4, Appli
130	42	43.8	91	5	US-10-732-923-19975	Sequence 19975, A	203	41	42.7	541	4	US-10-094-749-2778	Sequence 2778, Ap
131	42	43.8	102	4	US-10-424-599-211229	Sequence 211229, A	204	41	42.7	541	4	US-10-096-625-4	Sequence 4, Appli
132	42	43.8	107	4	US-10-732-923-19944	Sequence 39, Appl	205	41	42.7	639	4	US-10-282-122A-62843	Sequence 62843, A
133	42	43.8	107	4	US-10-732-923-19944	Sequence 44, Appl	206	41	42.7	639	4	US-10-282-122A-64694	Sequence 64694, A
134	42	43.8	109	4	US-10-424-599-224191	Sequence 224191, A	207	41	42.7	642	4	US-10-282-122A-63760	Sequence 63760, A
135	42	43.8	112	4	US-10-264-237-1922	Sequence 1922, Ap	208	41	42.7	747	4	US-10-425-115-222484	Sequence 222484, A
136	42	43.8	133	4	US-10-437-963-111922	Sequence 111922, A	209	41	42.7	774	4	US-10-280-962-44	Sequence 44, Appl
137	42	43.8	192	4	US-10-460-524-6	Sequence 6, Appli	210	41	42.7	774	4	US-10-298-680-44	Sequence 44, Appl
138	42	43.8	236	4	US-10-742-405-10	Sequence 10, Appl	211	41	42.7	774	4	US-10-702-400-54	Sequence 54, Appl
139	42	43.8	307	4	US-10-424-599-224191	Sequence 224191, A	212	41	42.7	774	5	US-10-805-650-54	Sequence 54, Appl
140	42	43.8	311	4	US-10-369-493-16772	Sequence 16772, A	213	41	42.7	889	5	US-10-737-450-80	Sequence 80, Appl
141	42	43.8	330	4	US-10-437-963-117534	Sequence 117534, A	214	41	42.7	971	5	US-10-450-763-45060	Sequence 45060, A
142	42	43.8	336	4	US-10-437-963-130681	Sequence 130681, A	215	41	42.7	972	4	US-10-360-522-61	Sequence 61, Appl
143	42	43.8	420	4	US-10-029-386-32104	Sequence 32104, A	216	41	42.7	992	4	US-10-360-522-56	Sequence 56, Appl
144	42	43.8	453	5	US-10-755-966-18	Sequence 18, Appl	217	41	42.7	992	4	US-10-360-522-60	Sequence 60, Appl
145	42	43.8	491	5	US-10-732-923-19974	Sequence 19, Appl	218	41	42.7	992	5	US-10-755-966-10	Sequence 10, Appl
146	42	43.8	491	5	US-10-732-923-19981	Sequence 19, Appl	219	41	42.7	1031	5	US-10-755-966-20	Sequence 20, Appl
147	42	43.8	491	5	US-10-732-923-19982	Sequence 19, Appl	220	41	42.7	1185	4	US-10-437-963-194044	Sequence 194044, A
148	42	43.8	492	5	US-10-732-923-19980	Sequence 19980, A	221	41	42.7	1233	4	US-10-369-493-3642	Sequence 3642, Ap
149	42	43.8	554	4	US-10-183-687-232	Sequence 232, App	222	41	42.7	1294	5	US-10-732-923-17000	Sequence 17000, A
150	42	43.8	615	4	US-10-291-172-356	Sequence 356, App	223	41	42.7	1315	4	US-10-437-963-128416	Sequence 128416, A
151	42	43.8	615	4	US-10-291-172-356	Sequence 356, App	224	40.5	42.2	228	4	US-10-767-701-51639	Sequence 51639, A
152	42	43.8	639	5	US-10-755-966-14	Sequence 14, Appl	225	40.5	42.2	367	4	US-10-425-114-58398	Sequence 58398, A
153	42	43.8	673	5	US-10-450-763-56458	Sequence 56458, A	226	40.5	42.2	428	4	US-10-425-114-53337	Sequence 53337, A
154	42	43.8	681	5	US-10-662-358-1	Sequence 1, Appli	227	40.5	42.2	428	4	US-10-425-114-59976	Sequence 59976, A
155	42	43.8	773	4	US-10-280-962-40	Sequence 40, Appl	228	40.5	42.2	430	4	US-10-437-963-129737	Sequence 129737, A
156	42	43.8	773	4	US-10-298-680-40	Sequence 40, Appl	229	40	41.7	78	4	US-10-424-599-200771	Sequence 200771, A
157	42	43.8	773	4	US-10-702-400-50	Sequence 50, Appl	230	40	41.7	82	5	US-10-450-763-35497	Sequence 35497, A
158	42	43.8	773	5	US-10-805-650-50	Sequence 50, Appl	231	40	41.7	83	4	US-10-437-963-179775	Sequence 179775, A
159	42	43.8	774	4	US-10-280-962-42	Sequence 42, Appl	232	40	41.7	87	4	US-10-425-115-311021	Sequence 311021, A
160	42	43.8	774	4	US-10-298-680-42	Sequence 42, Appl	233	40	41.7	103	4	US-10-282-122A-57503	Sequence 57503, A
161	42	43.8	774	4	US-10-702-400-52	Sequence 52, Appl	234	40	41.7	107	3	US-09-848-798-29	Sequence 29, Appl
162	42	43.8	774	5	US-10-805-650-52	Sequence 52, Appl	235	40	41.7	108	4	US-10-803-622-244	Sequence 244, App
163	42	43.8	775	4	US-10-280-962-38	Sequence 38, Appl	236	40	41.7	108	4	US-10-803-653-244	Sequence 244, App
164	42	43.8	775	4	US-10-298-680-38	Sequence 38, Appl	237	40	41.7	114	5	US-10-450-763-38824	Sequence 38824, A
165	42	43.8	775	4	US-10-702-400-48	Sequence 48, Appl	238	40	41.7	128	4	US-10-425-115-257389	Sequence 257389, A
166	42	43.8	775	5	US-10-805-650-48	Sequence 48, Appl	239	40	41.7	141	4	US-10-156-761-11934	Sequence 11934, A
167	42	43.8	779	4	US-10-126-757A-9	Sequence 9, Appli	240	40	41.7	163	4	US-10-424-599-173913	Sequence 173913, A
168	42	43.8	779	4	US-10-126-757A-9	Sequence 9, Appli	241	40	41.7	180	5	US-10-856-499-834	Sequence 834, App
169	42	43.8	904	4	US-10-437-963-198457	Sequence 9, Appli	242	40	41.7	240	5	US-10-893-671-98	Sequence 98, Appl
170	42	43.8	945	4	US-10-360-522-62	Sequence 62, Appl	243	40	41.7	252	4	US-10-767-701-40354	Sequence 40354, A
171	42	43.8	945	4	US-10-360-522-63	Sequence 63, Appl	244	40	41.7	263	4	US-10-104-047-2017	Sequence 2017, Ap
172	42	43.8	970	4	US-10-360-522-54	Sequence 54, Appl	245	40	41.7	285	4	US-10-425-114-46544	Sequence 46544, A
173	42	43.8	970	5	US-10-755-966-5	Sequence 5, Appli	246	40	41.7	292	4	US-10-425-114-48595	Sequence 48595, A



247	40	41.7	296	4	US-10-425-114-62044	Sequence 62044, A	320	39	40.6	411	4	US-10-425-114-43595	Sequence 43595, A
248	40	41.7	312	4	US-10-369-493-2598	Sequence 2598, A	321	39	40.6	411	4	US-10-425-115-335861	Sequence 335861, A
249	40	41.7	345	5	US-10-450-763-56494	Sequence 56494, A	322	39	40.6	415	4	US-10-425-114-72426	Sequence 72426, A
250	40	41.7	346	4	US-10-157-031-289	Sequence 289, App	323	39	40.6	418	4	US-10-706-691-18	Sequence 18, Appl
251	40	41.7	349	4	US-10-437-963-123444	Sequence 123444, A	324	39	40.6	440	4	US-10-114-270-10	Sequence 10, Appl
252	40	41.7	424	4	US-10-437-963-192370	Sequence 192370, A	325	39	40.6	442	4	US-10-258-106-8	Sequence 8, Appl
253	40	41.7	500	4	US-10-437-963-141185	Sequence 141185, A	326	39	40.6	444	4	US-10-425-115-223566	Sequence 223566, A
254	40	41.7	504	4	US-10-259-194A-174	Sequence 174, App	327	39	40.6	455	4	US-10-103-547A-2	Sequence 2, Appl
255	40	41.7	507	5	US-10-450-763-50922	Sequence 50922, A	328	39	40.6	455	4	US-10-114-270-12	Sequence 12, Appl
256	40	41.7	529	4	US-10-437-963-157883	Sequence 157883, A	329	39	40.6	455	4	US-10-433-794-8	Sequence 8, Appl
257	40	41.7	617	4	US-10-156-761-12250	Sequence 12250, A	330	39	40.6	455	4	US-10-655-490-2	Sequence 2, Appl
258	40	41.7	701	4	US-10-369-493-5420	Sequence 5420, App	331	39	40.6	476	4	US-10-424-599-283133	Sequence 283133, A
259	40	41.7	747	4	US-10-762-596-2	Sequence 2, Appl	332	39	40.6	477	3	US-09-994-288-2	Sequence 2, Appl
260	40	41.7	747	4	US-10-762-596-4	Sequence 4, Appl	333	39	40.6	488	4	US-10-074-978A-248	Sequence 248, App
261	40	41.7	831	4	US-10-369-493-7854	Sequence 7854, A	334	39	40.6	513	4	US-10-108-260A-2517	Sequence 2517, App
262	40	41.7	900	5	US-10-450-763-56493	Sequence 56493, A	335	39	40.6	534	4	US-10-108-260A-2517	Sequence 2517, App
263	40	41.7	1089	4	US-10-437-963-192363	Sequence 192363, A	336	39	40.6	537	3	US-10-282-122A-49400	Sequence 49400, A
264	40	41.7	1192	5	US-10-450-763-53577	Sequence 53577, A	337	39	40.6	548	4	US-09-486-734A-8	Sequence 8, Appl
265	40	41.7	1305	4	US-10-282-122A-76741	Sequence 76741, A	338	39	40.6	596	6	US-10-478-146-10	Sequence 10, Appl
266	39.5	41.1	98	4	US-10-437-963-132357	Sequence 132357, A	339	39	40.6	606	4	US-11-097-143-42057	Sequence 42057, A
267	39.5	41.1	244	4	US-10-437-963-117929	Sequence 117929, A	340	39	40.6	611	6	US-10-276-774-1383	Sequence 1383, App
268	39.5	41.1	299	4	US-10-369-493-15764	Sequence 15764, A	341	39	40.6	618	4	US-11-097-143-9399	Sequence 9399, App
269	39.5	41.1	299	4	US-10-369-493-16148	Sequence 16148, A	342	39	40.6	630	4	US-10-114-270-14	Sequence 14, Appl
270	39.5	41.1	303	4	US-10-369-493-15396	Sequence 15396, A	343	39	40.6	636	5	US-10-108-260A-3776	Sequence 3776, App
271	39.5	41.1	312	5	US-10-200-545-71	Sequence 71, Appl	344	39	40.6	637	3	US-10-450-763-57991	Sequence 57991, A
272	39.5	41.1	387	4	US-10-369-493-19896	Sequence 19896, A	345	39	40.6	646	4	US-10-282-122A-56091	Sequence 56091, A
273	39.5	41.1	543	5	US-10-732-923-1144	Sequence 1144, App	346	39	40.6	659	3	US-09-761-640-5	Sequence 5, Appl
274	39	40.6	34	4	US-10-425-115-184844	Sequence 184844, A	347	39	40.6	659	3	US-09-955-732-2	Sequence 2, Appl
275	39	40.6	70	4	US-10-106-698-4364	Sequence 4364, App	348	39	40.6	663	5	US-10-450-763-59562	Sequence 59562, A
276	39	40.6	77	4	US-10-425-115-272545	Sequence 272545, A	349	39	40.6	667	4	US-10-282-122A-42875	Sequence 42875, A
277	39	40.6	79	4	US-10-425-115-358291	Sequence 358291, A	350	39	40.6	670	3	US-09-833-245-1304	Sequence 1304, App
278	39	40.6	90	4	US-10-437-963-186699	Sequence 186699, A	351	39	40.6	671	5	US-10-450-763-48139	Sequence 48139, A
279	39	40.6	90	4	US-10-425-115-205072	Sequence 205072, A	352	39	40.6	675	4	US-10-108-260A-3326	Sequence 3326, App
280	39	40.6	91	4	US-10-425-115-269515	Sequence 269515, A	353	39	40.6	687	4	US-10-369-493-368	Sequence 368, App
281	39	40.6	100	4	US-10-424-599-233944	Sequence 233944, A	354	39	40.6	692	4	US-10-369-493-6712	Sequence 6712, App
282	39	40.6	102	4	US-10-425-115-192242	Sequence 192242, A	355	39	40.6	711	3	US-09-738-626-4854	Sequence 4854, App
283	39	40.6	107	4	US-10-723-434-36	Sequence 36, Appl	356	39	40.6	731	4	US-10-781-014-362	Sequence 362, App
284	39	40.6	108	5	US-10-706-689-19	Sequence 19, Appl	357	39	40.6	739	4	US-10-425-114-54476	Sequence 54476, A
285	39	40.6	108	5	US-10-988-360-19	Sequence 19, Appl	358	39	40.6	755	4	US-10-156-761-9852	Sequence 9852, App
286	39	40.6	126	4	US-10-767-701-60623	Sequence 60623, A	359	39	40.6	774	3	US-09-803-165-34	Sequence 34, Appl
287	39	40.6	134	4	US-10-425-115-271453	Sequence 271453, A	360	39	40.6	779	4	US-10-363-676-11	Sequence 11, Appl
288	39	40.6	143	4	US-10-424-599-202808	Sequence 202808, A	361	39	40.6	797	5	US-10-450-763-59892	Sequence 59892, A
289	39	40.6	143	4	US-10-425-114-37336	Sequence 37336, A	362	39	40.6	805	4	US-10-429-160-60	Sequence 60, Appl
290	39	40.6	143	3	US-09-867-550-66	Sequence 66, Appl	363	39	40.6	805	5	US-10-719-993-546	Sequence 546, App
291	39	40.6	150	4	US-10-425-115-319178	Sequence 319178, A	364	39	40.6	865	5	US-10-719-993-547	Sequence 547, App
292	39	40.6	150	5	US-10-450-763-38974	Sequence 38974, A	365	39	40.6	959	6	US-11-097-143-22896	Sequence 22896, A
293	39	40.6	161	4	US-10-108-260A-3409	Sequence 3409, App	366	39	40.6	984	4	US-10-369-493-8357	Sequence 8357, App
294	39	40.6	169	4	US-10-425-115-246213	Sequence 246213, A	367	39	40.6	1028	4	US-10-764-425-171	Sequence 171, App
295	39	40.6	182	4	US-10-425-115-246213	Sequence 246213, A	368	39	40.6	1036	5	US-10-819-386A-12	Sequence 12, Appl
296	39	40.6	190	4	US-10-425-115-289998	Sequence 289998, A	369	39	40.6	1047	4	US-10-437-963-138135	Sequence 138135, A
297	39	40.6	207	4	US-10-282-122A-50000	Sequence 50000, A	370	39	40.6	1254	4	US-10-369-493-22131	Sequence 22131, A
298	39	40.6	215	4	US-10-462-040A-2	Sequence 2, Appl	371	39	40.6	1265	3	US-09-862-027-80	Sequence 80, Appl
299	39	40.6	227	5	US-10-739-930-7782	Sequence 7782, App	372	39	40.6	1255	5	US-10-989-228-80	Sequence 80, Appl
300	39	40.6	229	4	US-10-425-115-289999	Sequence 289999, A	373	39	40.6	1433	4	US-10-788-792-232	Sequence 232, App
301	39	40.6	240	4	US-10-083-357-1279	Sequence 1279, App	374	39	40.6	1611	4	US-10-425-114-54372	Sequence 54372, A
302	39	40.6	244	4	US-10-425-114-64967	Sequence 64967, A	375	39	40.6	1647	5	US-10-450-763-41109	Sequence 41109, A
303	39	40.6	279	5	US-10-474-955-98	Sequence 98, Appl	376	39	40.6	1720	4	US-10-433-794-9	Sequence 9, Appl
304	39	40.6	279	5	US-10-474-955-99	Sequence 99, Appl	377	39	40.6	1734	3	US-09-862-027-81	Sequence 81, Appl
305	39	40.6	279	5	US-10-474-955-100	Sequence 100, Appl	378	39	40.6	1734	4	US-10-377-035-14	Sequence 14, Appl
306	39	40.6	279	5	US-10-474-955-100	Sequence 100, App	379	39	40.6	1734	4	US-10-042-865-82	Sequence 82, Appl
307	39	40.6	284	4	US-10-425-114-50177	Sequence 50177, A	380	39	40.6	1734	5	US-10-737-450-136	Sequence 136, App
308	39	40.6	291	4	US-10-425-115-206991	Sequence 206991, A	381	39	40.6	1734	5	US-10-989-228-81	Sequence 81, Appl
309	39	40.6	294	4	US-10-437-963-179834	Sequence 179834, A	382	39	40.6	1798	4	US-10-311-034-23	Sequence 23, Appl
310	39	40.6	298	5	US-10-739-930-9621	Sequence 9621, App	383	39	40.6	1798	4	US-10-618-941-70	Sequence 70, Appl
311	39	40.6	318	3	US-09-954-342-26	Sequence 26, Appl	384	39	40.6	1805	4	US-10-425-114-54369	Sequence 54369, A
312	39	40.6	319	4	US-10-282-122A-42703	Sequence 42703, A	385	39	40.6	1829	4	US-10-435-766-13	Sequence 13, Appl
313	39	40.6	338	3	US-09-960-217-35	Sequence 35, Appl	386	39	40.6	1829	5	US-10-734-563-84	Sequence 84, Appl
314	39	40.6	353	4	US-10-437-963-186466	Sequence 186466, A	387	39	40.6	1829	5	US-10-853-973A-13	Sequence 13, Appl
315	39	40.6	357	4	US-10-288-930-93	Sequence 93, Appl	388	38.5	40.1	41	4	US-10-425-115-363785	Sequence 363785, A
316	39	40.6	374	4	US-10-664-421-96	Sequence 96, Appl	389	38.5	40.1	67	4	US-10-029-386-29229	Sequence 29229, A
317	39	40.6	374	5	US-10-941-635-96	Sequence 96, Appl	390	38.5	40.1	110	4	US-10-424-599-147764	Sequence 147764, A
318	39	40.6	378	4	US-10-450-763-45654	Sequence 45654, A	391	38.5	40.1	118	4	US-10-425-115-347459	Sequence 347459, A
319	39	40.6	384	4	US-10-425-114-70813	Sequence 70813, A	392	38.5	40.1	170	4	US-10-425-115-325981	Sequence 325981, A

393	38.5	40.1	367	5	US-10-739-930-7545	Sequence 7545, Ap	466	38	39.6	198	4	US-10-282-122A-59217	Sequence 59217, A
394	38.5	40.1	390	4	US-10-214-446-10	Sequence 10, Appl	467	38	39.6	199	4	US-10-437-963-157706	Sequence 157706,
395	38.5	40.1	425	5	US-10-450-763-42841	Sequence 42841, A	468	38	39.6	208	4	US-10-437-963-107730	Sequence 107730,
396	38.5	40.1	922	5	US-10-844-716-2	Sequence 2, Appl	469	38	39.6	209	4	US-10-767-701-38544	Sequence 38544, A
397	38.5	40.1	1330	4	US-10-282-122A-54113	Sequence 54113, A	470	38	39.6	220	3	US-09-220-920-26	Sequence 26, Appl
398	38.5	40.1	1485	4	US-10-262-794A-32	Sequence 32, Appl	471	38	39.6	220	3	US-09-804-615-9	Sequence 9, Appl
399	38.5	40.1	1653	3	US-09-741-669-402	Sequence 402, App	472	38	39.6	220	4	US-10-001-054-56	Sequence 56, Appl
400	38	39.6	20	5	US-10-474-955-21	Sequence 21, Appl	473	38	39.6	220	4	US-10-223-085-318	Sequence 318, App
401	38	39.6	34	4	US-10-425-115-232801	Sequence 232801,	474	38	39.6	220	4	US-10-223-084-318	Sequence 318, App
402	38	39.6	39	3	US-09-220-920-48	Sequence 48, Appl	475	38	39.6	220	4	US-10-223-088-318	Sequence 318, App
403	38	39.6	39	3	US-10-864-891-24	Sequence 24, Appl	476	38	39.6	220	4	US-10-223-090-318	Sequence 318, App
404	38	39.6	39	5	US-10-957-221-28	Sequence 28, Appl	477	38	39.6	220	4	US-10-223-087-318	Sequence 318, App
405	38	39.6	69	4	US-10-437-963-153588	Sequence 153588,	478	38	39.6	220	4	US-10-223-083-318	Sequence 318, App
406	38	39.6	81	5	US-10-732-923-19884	Sequence 19884, A	479	38	39.6	220	4	US-10-223-089-318	Sequence 318, App
407	38	39.6	84	4	US-10-029-386-29345	Sequence 29345, A	480	38	39.6	220	4	US-10-210-951-62	Sequence 62, Appl
408	38	39.6	84	5	US-10-450-763-45446	Sequence 45446, A	481	38	39.6	220	4	US-10-211-884-62	Sequence 62, Appl
409	38	39.6	92	4	US-10-282-122A-65540	Sequence 65540, A	482	38	39.6	220	4	US-10-223-081-318	Sequence 318, App
410	38	39.6	92	4	US-10-282-122A-65639	Sequence 65639, A	483	38	39.6	220	4	US-10-223-082-318	Sequence 318, App
411	38	39.6	92	4	US-10-424-599-207663	Sequence 207663,	484	38	39.6	220	4	US-10-211-858-62	Sequence 62, Appl
412	38	39.6	97	4	US-10-425-115-292097	Sequence 292097,	485	38	39.6	220	4	US-10-305-654-318	Sequence 318, App
413	38	39.6	100	4	US-10-437-963-114573	Sequence 114573,	486	38	39.6	220	4	US-10-285-027-402	Sequence 402, App
414	38	39.6	102	5	US-10-450-763-45444	Sequence 45444, A	487	38	39.6	220	4	US-10-295-027-404	Sequence 404, App
415	38	39.6	104	4	US-10-282-122A-60745	Sequence 60745, A	488	38	39.6	220	4	US-10-081-056-318	Sequence 318, App
416	38	39.6	105	4	US-10-424-599-165066	Sequence 165066,	489	38	39.6	220	4	US-10-669-853-2	Sequence 2, Appl
417	38	39.6	106	4	US-10-425-115-219841	Sequence 219841,	490	38	39.6	220	4	US-10-661-984A-9	Sequence 9, Appl
418	38	39.6	107	3	US-09-220-920-52	Sequence 52, Appl	491	38	39.6	220	5	US-10-806-793-9	Sequence 9, Appl
419	38	39.6	107	4	US-10-251-085B-127	Sequence 127, App	492	38	39.6	220	5	US-10-806-793-35	Sequence 35, Appl
420	38	39.6	107	4	US-10-737-252-127	Sequence 127, App	493	38	39.6	220	5	US-10-451-567-3	Sequence 3, Appl
421	38	39.6	113	4	US-10-424-599-208203	Sequence 208203,	494	38	39.6	220	5	US-10-864-891-10	Sequence 10, Appl
422	38	39.6	113	4	US-10-437-963-183059	Sequence 183059,	495	38	39.6	220	5	US-10-356-264A-5	Sequence 5, Appl
423	38	39.6	117	4	US-10-425-115-17684	Sequence 17684,	496	38	39.6	220	5	US-10-957-221-24	Sequence 24, Appl
424	38	39.6	118	5	US-10-491-653-148	Sequence 148, App	497	38	39.6	220	5	US-10-794-801-3	Sequence 3, Appl
425	38	39.6	127	4	US-10-425-115-317685	Sequence 317685,	498	38	39.6	220	6	US-11-074-498-10	Sequence 10, Appl
426	38	39.6	128	4	US-10-437-963-103853	Sequence 103853,	499	38	39.6	222	4	US-10-437-963-157707	Sequence 157707,
427	38	39.6	129	4	US-10-425-115-248772	Sequence 248772,	500	38	39.6	224	4	US-10-282-122A-50537	Sequence 50537, A
428	38	39.6	130	4	US-10-425-115-350417	Sequence 350417,	501	38	39.6	225	4	US-10-425-115-307019	Sequence 307019,
429	38	39.6	131	4	US-10-425-115-197312	Sequence 197312,	502	38	39.6	227	5	US-10-739-930-8496	Sequence 8496, Ap
430	38	39.6	131	4	US-10-425-115-197316	Sequence 197316,	503	38	39.6	239	5	US-10-732-923-5681	Sequence 5681, Ap
431	38	39.6	135	4	US-10-767-701-54703	Sequence 54703, A	504	38	39.6	242	4	US-10-424-599-24397	Sequence 243497,
432	38	39.6	136	3	US-09-925-300-1098	Sequence 1098, Ap	505	38	39.6	261	4	US-10-425-115-246261	Sequence 246261,
433	38	39.6	139	5	US-10-864-891-30	Sequence 30, Appl	506	38	39.6	262	4	US-10-000-954-8	Sequence 8, Appl
434	38	39.6	139	4	US-10-263-828-120	Sequence 120, App	507	38	39.6	271	4	US-10-425-115-232288	Sequence 232288,
435	38	39.6	139	4	US-10-425-115-256702	Sequence 256702,	508	38	39.6	278	4	US-10-437-963-185910	Sequence 185910,
436	38	39.6	141	4	US-10-425-114-48129	Sequence 48129, A	509	38	39.6	279	4	US-10-104-047-3382	Sequence 3382, Ap
437	38	39.6	141	4	US-10-425-114-59320	Sequence 59320, A	510	38	39.6	280	5	US-10-732-923-18987	Sequence 18987, A
438	38	39.6	141	4	US-10-425-114-61407	Sequence 61407, A	511	38	39.6	298	3	US-09-815-242-5142	Sequence 5142, Ap
439	38	39.6	141	4	US-10-425-114-61749	Sequence 61749, A	512	38	39.6	298	4	US-10-282-122A-43536	Sequence 43536, A
440	38	39.6	141	4	US-10-425-114-63509	Sequence 63509, A	513	38	39.6	308	5	US-10-287-226-356	Sequence 356, App
441	38	39.6	141	4	US-10-425-114-67251	Sequence 67251, A	514	38	39.6	308	5	US-10-732-923-19126	Sequence 19126, A
442	38	39.6	141	5	US-10-864-891-29	Sequence 29, Appl	515	38	39.6	308	5	US-10-732-923-19128	Sequence 19128, A
443	38	39.6	143	5	US-10-864-891-28	Sequence 28, Appl	516	38	39.6	308	5	US-10-732-923-19129	Sequence 19129, A
444	38	39.6	143	5	US-10-957-221-8	Sequence 8, Appl	517	38	39.6	308	5	US-10-732-923-19130	Sequence 19130, A
445	38	39.6	145	5	US-10-864-891-27	Sequence 27, Appl	518	38	39.6	316	4	US-10-767-701-45123	Sequence 45123, A
446	38	39.6	146	4	US-10-425-114-50621	Sequence 50621, A	519	38	39.6	324	4	US-10-156-761-12339	Sequence 12339, A
447	38	39.6	147	4	US-10-425-114-46087	Sequence 46087, A	520	38	39.6	324	5	US-10-732-923-9983	Sequence 9983, Ap
448	38	39.6	148	4	US-10-425-114-47016	Sequence 47016, A	521	38	39.6	326	4	US-10-369-493-12788	Sequence 12788, A
449	38	39.6	150	4	US-10-425-114-63322	Sequence 63322, A	522	38	39.6	327	4	US-10-437-963-123590	Sequence 123590,
450	38	39.6	152	5	US-10-864-891-26	Sequence 26, Appl	523	38	39.6	329	3	US-09-924-841-5	Sequence 5, Appl
451	38	39.6	153	4	US-10-425-115-201149	Sequence 201149,	524	38	39.6	329	3	US-09-924-841-9	Sequence 9, Appl
452	38	39.6	153	4	US-10-425-115-216305	Sequence 216305,	525	38	39.6	330	4	US-10-767-701-40033	Sequence 40033, A
453	38	39.6	156	4	US-10-425-114-59890	Sequence 59890, A	526	38	39.6	332	4	US-10-425-115-216640	Sequence 216640,
454	38	39.6	156	4	US-10-425-114-62502	Sequence 62502, A	527	38	39.6	334	4	US-10-424-599-190993	Sequence 190993,
455	38	39.6	159	4	US-10-425-115-317687	Sequence 317687,	528	38	39.6	335	4	US-10-425-114-47803	Sequence 47803, A
456	38	39.6	160	4	US-10-425-114-40414	Sequence 40414, A	529	38	39.6	341	4	US-10-425-114-64303	Sequence 64303, A
457	38	39.6	161	4	US-10-425-115-228644	Sequence 228644,	530	38	39.6	342	4	US-10-437-963-143458	Sequence 143458,
458	38	39.6	171	4	US-10-767-701-48906	Sequence 48906, A	531	38	39.6	347	4	US-10-437-963-142844	Sequence 142844,
459	38	39.6	177	4	US-10-425-114-60044	Sequence 60044, A	532	38	39.6	349	4	US-10-115-571A-33	Sequence 33, Appl
460	38	39.6	179	5	US-10-864-891-25	Sequence 25, Appl	533	38	39.6	366	4	US-10-425-115-267141	Sequence 267141,
461	38	39.6	181	4	US-10-425-114-54895	Sequence 54895, A	534	38	39.6	374	4	US-10-282-122A-60826	Sequence 60826, A
462	38	39.6	185	4	US-10-425-115-296754	Sequence 296754,	535	38	39.6	382	4	US-10-128-714-3222	Sequence 3222, Ap
463	38	39.6	190	4	US-10-437-963-194870	Sequence 194870,	536	38	39.6	383	5	US-10-732-923-4805	Sequence 4805, Ap
464	38	39.6	191	5	US-10-732-923-18992	Sequence 18992, A	537	38	39.6	383	5	US-10-732-923-4806	Sequence 4806, Ap
465	38	39.6	196	4	US-10-425-115-312485	Sequence 312485,	538	38	39.6	404	4	US-10-425-115-211677	Sequence 211677,

539	38	39.6	404	5	US-10-739-330-9718	Sequence 9718, Ap	612	38	39.6	775	4	US-10-298-680-32	Sequence 32, Appl
540	38	39.6	409	3	US-09-815-242-11699	Sequence 11699, A	613	38	39.6	775	4	US-10-298-680-35	Sequence 35, Appl
541	38	39.6	410	6	US-11-097-143-15597	Sequence 15597, A	614	38	39.6	775	4	US-10-702-400-42	Sequence 42, Appl
542	38	39.6	441	4	US-10-369-493-349	Sequence 349, App	615	38	39.6	775	4	US-10-702-400-45	Sequence 45, Appl
543	38	39.6	451	4	US-10-282-122A-55907	Sequence 55907, A	616	38	39.6	775	5	US-10-805-650-42	Sequence 42, Appl
544	38	39.6	461	4	US-10-425-114-69472	Sequence 69472, A	617	38	39.6	775	5	US-10-805-650-45	Sequence 45, Appl
545	38	39.6	478	4	US-10-366-345-32	Sequence 32, Appl	618	38	39.6	775	5	US-10-734-563-91	Sequence 91, Appl
546	38	39.6	478	4	US-10-295-027-656	Sequence 656, App	619	38	39.6	775	5	US-10-734-563-101	Sequence 101, App
547	38	39.6	478	5	US-10-723-980-2346	Sequence 2346, Ap	620	38	39.6	775	5	US-10-853-973A-15	Sequence 15, Appl
548	38	39.6	479	3	US-09-813-386-26	Sequence 26, Appl	621	38	39.6	776	4	US-10-280-962-45	Sequence 45, Appl
549	38	39.6	479	5	US-10-826-324-26	Sequence 26, Appl	622	38	39.6	776	4	US-10-280-962-45	Sequence 45, Appl
550	38	39.6	497	3	US-09-764-893-125	Sequence 125, App	623	38	39.6	776	4	US-10-398-680-45	Sequence 45, Appl
551	38	39.6	497	3	US-08-764-881-157	Sequence 157, App	624	38	39.6	776	5	US-10-702-400-55	Sequence 55, Appl
552	38	39.6	497	3	US-09-764-853-816	Sequence 816, App	625	38	39.6	828	4	US-10-805-650-55	Sequence 55, Appl
553	38	39.6	497	3	US-09-764-898-276	Sequence 276, App	626	38	39.6	828	4	US-10-369-493-901	Sequence 901, App
554	38	39.6	497	3	US-09-764-881-157	Sequence 157, App	627	38	39.6	828	4	US-10-282-122A-67347	Sequence 67347, A
555	38	39.6	497	3	US-09-764-875-1154	Sequence 1154, App	628	38	39.6	829	4	US-10-369-493-14754	Sequence 14754, A
556	38	39.6	497	4	US-10-073-965-125	Sequence 125, App	629	38	39.6	834	4	US-10-369-493-11481	Sequence 11481, A
557	38	39.6	497	4	US-10-242-747-157	Sequence 157, App	630	38	39.6	834	4	US-10-369-493-15233	Sequence 15233, A
558	38	39.6	553	4	US-10-238-075-802	Sequence 802, App	631	38	39.6	886	4	US-10-437-963-141001	Sequence 141001, A
559	38	39.6	556	5	US-10-450-763-37737	Sequence 37737, A	632	38	39.6	934	4	US-10-282-122A-68707	Sequence 68707, A
560	38	39.6	577	4	US-10-369-493-23464	Sequence 23464, A	633	38	39.6	935	5	US-10-450-763-40531	Sequence 40531, A
561	38	39.6	596	4	US-10-229-346-7	Sequence 7, Appli	634	38	39.6	944	5	US-10-972-024-270	Sequence 270, App
562	38	39.6	596	4	US-10-487-846-7	Sequence 7, Appli	635	38	39.6	976	4	US-10-369-493-16784	Sequence 16784, A
563	38	39.6	596	4	US-10-229-346-4	Sequence 4, Appli	636	38	39.6	1000	5	US-10-647-268-8	Sequence 8, Appli
564	38	39.6	597	4	US-10-229-346-13	Sequence 13, Appl	637	38	39.6	1031	4	US-10-087-192-93	Sequence 93, Appl
565	38	39.6	597	4	US-10-229-346-21	Sequence 21, Appl	638	38	39.6	1039	4	US-10-087-192-96	Sequence 96, Appl
566	38	39.6	597	4	US-10-102-469-16	Sequence 16, Appl	639	38	39.6	1044	3	US-09-893-371-1	Sequence 1, Appli
567	38	39.6	597	4	US-10-487-846-4	Sequence 4, Appli	640	38	39.6	1044	3	US-10-287-216-1	Sequence 1, Appli
568	38	39.6	597	4	US-10-487-846-13	Sequence 13, Appl	641	38	39.6	1050	5	US-10-450-763-40530	Sequence 40530, A
569	38	39.6	597	4	US-10-487-846-21	Sequence 21, Appl	642	38	39.6	1139	3	US-09-764-893-102	Sequence 102, App
570	38	39.6	598	4	US-10-229-346-9	Sequence 9, Appli	643	38	39.6	1139	3	US-09-764-881-99	Sequence 99, Appl
571	38	39.6	598	4	US-10-487-846-9	Sequence 9, Appli	644	38	39.6	1139	3	US-09-764-853-634	Sequence 634, App
572	38	39.6	600	4	US-10-229-346-17	Sequence 17, Appl	645	38	39.6	1139	3	US-09-764-898-204	Sequence 204, App
573	38	39.6	600	4	US-10-487-846-17	Sequence 17, Appl	646	38	39.6	1139	3	US-09-764-881-99	Sequence 99, Appl
574	38	39.6	601	4	US-10-229-346-15	Sequence 15, Appl	647	38	39.6	1139	3	US-09-764-875-873	Sequence 873, App
575	38	39.6	601	4	US-10-487-846-15	Sequence 15, Appl	648	38	39.6	1139	4	US-10-073-865-102	Sequence 102, App
576	38	39.6	602	4	US-10-229-346-19	Sequence 19, Appl	649	38	39.6	1251	4	US-10-742-747-99	Sequence 99, Appl
577	38	39.6	602	4	US-10-487-846-19	Sequence 19, Appl	650	38	39.6	1251	4	US-10-311-623-12	Sequence 12, Appl
578	38	39.6	603	4	US-10-156-761-13963	Sequence 13963, A	651	38	39.6	1276	4	US-10-156-761-10509	Sequence 10509, A
579	38	39.6	606	4	US-10-229-346-11	Sequence 11, Appl	652	38	39.6	1316	6	US-10-870-507A-2	Sequence 2, Appli
580	38	39.6	606	4	US-10-487-846-11	Sequence 11, Appl	653	38	39.6	1316	6	US-11-097-143-5979	Sequence 5979, Ap
581	38	39.6	613	6	US-11-097-143-42687	Sequence 42687, A	654	38	39.6	1422	5	US-10-450-763-32704	Sequence 32704, A
582	38	39.6	618	4	US-10-425-115-278201	Sequence 278201, A	655	38	39.6	1487	5	US-10-626-832-77	Sequence 77, Appl
583	38	39.6	627	4	US-10-156-761-12875	Sequence 12875, A	656	38	39.6	1487	5	US-10-626-832-100	Sequence 100, App
584	38	39.6	644	3	US-09-943-692-2	Sequence 2, Appli	657	38	39.6	1498	4	US-10-017-161-1658	Sequence 1658, Ap
585	38	39.6	644	4	US-10-222-441-2	Sequence 2, Appli	658	38	39.6	1498	4	US-10-292-798-1322	Sequence 1322, Ap
586	38	39.6	644	4	US-10-229-346-2	Sequence 2, Appli	659	38	39.6	1568	4	US-10-263-929-117	Sequence 117, App
587	38	39.6	644	4	US-10-487-846-2	Sequence 2, Appli	660	38	39.6	2050	3	US-09-898-570-22	Sequence 22, Appl
588	38	39.6	651	4	US-10-614-076-52	Sequence 52, Appl	661	38	39.6	2050	3	US-09-839-446-22	Sequence 22, Appl
589	38	39.6	651	4	US-10-614-076-56	Sequence 56, Appl	662	38	39.6	3557	4	US-10-295-027-430	Sequence 430, App
590	38	39.6	651	4	US-10-614-076-58	Sequence 58, Appl	663	38	39.6	3557	4	US-10-295-027-1297	Sequence 1297, Ap
591	38	39.6	652	4	US-10-614-076-113	Sequence 113, App	664	38	39.6	3568	4	US-10-028-248A-8	Sequence 8, Appli
592	38	39.6	652	4	US-10-782-020-9	Sequence 9, Appli	665	38	39.6	3568	4	US-10-107-782-8	Sequence 8, Appli
593	38	39.6	652	4	US-10-782-141-10	Sequence 10, Appl	666	38	39.6	3570	4	US-10-028-248A-6	Sequence 6, Appli
594	38	39.6	652	4	US-10-782-096-11	Sequence 11, Appl	667	38	39.6	3570	4	US-10-107-782-6	Sequence 6, Appli
595	38	39.6	652	4	US-10-782-570-8	Sequence 8, Appli	668	38	39.6	3571	3	US-09-911-842-2	Sequence 2, Appli
596	38	39.6	652	5	US-10-783-417-6	Sequence 6, Appli	669	38	39.6	3571	4	US-10-150-821-2	Sequence 2, Appli
597	38	39.6	652	5	US-10-781-379-12	Sequence 12, Appl	670	38	39.6	3571	4	US-10-603-283-2	Sequence 2, Appli
598	38	39.6	652	5	US-10-926-819-13	Sequence 13, Appl	671	38	39.6	4019	3	US-09-738-973-425	Sequence 425, App
599	38	39.6	653	4	US-10-282-122A-60290	Sequence 60290, A	672	38	39.6	4019	3	US-09-854-133-425	Sequence 425, App
600	38	39.6	665	4	US-10-128-714-8222	Sequence 8222, Ap	673	37.5	39.1	73	3	US-10-144-649A-425	Sequence 425, App
601	38	39.6	669	4	US-10-282-122A-55934	Sequence 55934, A	674	37.5	39.1	77	3	US-10-424-599-147370	Sequence 147370, A
602	38	39.6	680	4	US-10-425-114-57368	Sequence 57368, A	675	37.5	39.1	139	4	US-10-864-761-39651	Sequence 39651, A
603	38	39.6	684	4	US-10-104-047-2418	Sequence 2418, Ap	676	37.5	39.1	139	4	US-10-425-115-328458	Sequence 328458, A
604	38	39.6	720	4	US-10-369-493-12449	Sequence 12449, A	677	37.5	39.1	156	3	US-09-764-864-1208	Sequence 1208, Ap
605	38	39.6	730	5	US-10-276-825-4	Sequence 4, Appli	678	37.5	39.1	244	5	US-10-450-763-34497	Sequence 34497, A
606	38	39.6	749	4	US-10-211-962-97	Sequence 97, Appl	679	37.5	39.1	245	4	US-10-425-115-206521	Sequence 206521, A
607	38	39.6	750	4	US-10-437-963-143294	Sequence 143294, A	680	37.5	39.1	253	4	US-10-156-761-11233	Sequence 11233, A
608	38	39.6	775	4	US-10-435-766-15	Sequence 15, Appl	681	37.5	39.1	293	4	US-10-437-963-180454	Sequence 180454, A
609	38	39.6	775	4	US-10-401-403-193	Sequence 193, App	682	37.5	39.1	294	4	US-10-264-237-2003	Sequence 2003, Ap
610	38	39.6	775	4	US-10-280-962-32	Sequence 32, Appl	683	37.5	39.1	476	4	US-10-104-047-3040	Sequence 3040, Ap
611	38	39.6	775	4	US-10-280-962-35	Sequence 35, Appl	684	37.5	39.1	504	4	US-10-104-047-2889	Sequence 2889, Ap
										525	4	US-10-163-198-2	Sequence 2, Appli

685	37.5	39.1	558	3	US-09-876-889-348	Sequence 348, App	758	37	38.5	184	4	US-10-424-599-162183	Sequence 162183,
686	37.5	39.1	620	4	US-10-437-963-160069	Sequence 160069,	759	37	38.5	186	4	US-10-437-963-160259	Sequence 160259,
687	37.5	39.1	625	4	US-10-259-194A-376	Sequence 376, App	760	37	38.5	192	4	US-10-437-963-189186	Sequence 189186,
688	37.5	39.1	684	4	US-09-493-3379	Sequence 3379, App	761	37	38.5	195	4	US-10-296-115-893	Sequence 893, App
689	37.5	39.1	770	3	US-09-784-316-5	Sequence 5, Appli	762	37	38.5	202	6	US-11-097-143-10083	Sequence 10083, A
690	37.5	39.1	770	4	US-10-229-124-5	Sequence 5, Appli	763	37	38.5	204	5	US-10-739-930-9682	Sequence 9682, Ap
691	37.5	39.1	770	4	US-10-778-301-5	Sequence 5, Appli	764	37	38.5	211	4	US-10-108-260A-4710	Sequence 4710, Ap
692	37.5	39.1	855	5	US-10-481-032A-294	Sequence 294, App	765	37	38.5	212	4	US-10-437-963-189184	Sequence 189184,
693	37.5	39.1	886	3	US-09-784-316-4	Sequence 4, Appli	766	37	38.5	224	4	US-10-029-386-33710	Sequence 33710, A
694	37.5	39.1	886	4	US-10-229-124-4	Sequence 4, Appli	767	37	38.5	224	4	US-10-437-963-110542	Sequence 110542,
695	37.5	39.1	886	4	US-10-778-301-4	Sequence 4, Appli	768	37	38.5	225	5	US-10-732-523-5731	Sequence 5731, Ap
696	37.5	39.1	886	4	US-10-778-301-4	Sequence 4, Appli	769	37	38.5	226	4	US-10-029-386-32061	Sequence 32061, A
697	37.5	39.1	887	3	US-09-771-404-2	Sequence 2, Appli	770	37	38.5	226	4	US-10-425-114-51827	Sequence 51827, A
698	37.5	39.1	887	4	US-10-263-316-2	Sequence 2, Appli	771	37	38.5	232	4	US-10-425-115-225109	Sequence 225109,
699	37.5	39.1	1067	4	US-10-437-963-175718	Sequence 175718,	772	37	38.5	234	4	US-10-320-797-3040	Sequence 3040, Ap
700	37	38.5	34	4	US-10-120-885A-21	Sequence 21, Appl	773	37	38.5	239	4	US-10-425-115-238116	Sequence 238116,
701	37	38.5	35	4	US-10-425-115-353079	Sequence 353079,	774	37	38.5	240	4	US-10-156-761-9685	Sequence 9685, Ap
702	37	38.5	62	4	US-10-437-963-195640	Sequence 195640,	775	37	38.5	244	4	US-10-425-114-50501	Sequence 50501, A
703	37	38.5	65	4	US-10-425-115-223490	Sequence 223490,	776	37	38.5	245	4	US-10-425-114-49293	Sequence 49293, A
704	37	38.5	69	4	US-10-425-115-289742	Sequence 289742,	777	37	38.5	248	4	US-10-437-963-139365	Sequence 139365,
705	37	38.5	72	4	US-10-437-963-113104	Sequence 113104,	778	37	38.5	250	4	US-10-437-963-110546	Sequence 110546,
706	37	38.5	74	4	US-10-425-114-40456	Sequence 40456, A	779	37	38.5	253	4	US-10-291-172-732	Sequence 732, App
707	37	38.5	74	4	US-10-425-114-40456	Sequence 40456, A	780	37	38.5	253	4	US-10-221-278-732	Sequence 732, App
708	37	38.5	75	3	US-09-864-408A-132	Sequence 132, App	781	37	38.5	266	4	US-10-369-493-4469	Sequence 4469, Ap
709	37	38.5	80	4	US-10-424-599-271695	Sequence 271695,	782	37	38.5	266	4	US-10-369-493-7229	Sequence 7229, Ap
710	37	38.5	82	3	US-09-867-550-1704	Sequence 1704, Ap	783	37	38.5	272	5	US-10-450-763-30476	Sequence 30476, A
711	37	38.5	83	4	US-10-425-115-268921	Sequence 268921,	784	37	38.5	272	5	US-10-450-763-30476	Sequence 30476, A
712	37	38.5	85	4	US-10-437-963-144130	Sequence 144130,	785	37	38.5	276	4	US-10-156-761-12124	Sequence 12124, A
713	37	38.5	86	4	US-10-425-115-342319	Sequence 342319,	786	37	38.5	277	6	US-11-097-143-37368	Sequence 37368, A
714	37	38.5	87	4	US-10-425-115-222129	Sequence 222129,	787	37	38.5	280	4	US-10-425-115-204744	Sequence 204744,
715	37	38.5	90	5	US-10-767-701-31572	Sequence 31572, A	788	37	38.5	282	5	US-10-474-955-101	Sequence 101, App
716	37	38.5	93	4	US-10-450-763-57945	Sequence 57945, A	789	37	38.5	294	4	US-10-225-066A-964	Sequence 964, App
717	37	38.5	94	4	US-10-425-114-57563	Sequence 57563, A	790	37	38.5	294	4	US-10-374-780A-424	Sequence 424, App
718	37	38.5	94	4	US-10-767-701-37671	Sequence 37671, A	791	37	38.5	294	4	US-10-225-066A-964	Sequence 964, App
719	37	38.5	95	4	US-10-437-963-197628	Sequence 197628,	792	37	38.5	300	4	US-10-424-599-159337	Sequence 159337,
720	37	38.5	96	4	US-10-263-959-1204	Sequence 1204, Ap	793	37	38.5	312	4	US-10-282-122A-47447	Sequence 47447, A
721	37	38.5	96	4	US-10-425-115-278991	Sequence 278991,	794	37	38.5	315	4	US-10-425-114-53528	Sequence 53528, A
722	37	38.5	100	4	US-10-425-114-56915	Sequence 56915, A	795	37	38.5	316	5	US-10-450-763-58641	Sequence 58641, A
723	37	38.5	104	5	US-10-450-763-45452	Sequence 45452, A	796	37	38.5	323	4	US-10-425-114-67908	Sequence 67908, A
724	37	38.5	105	4	US-10-425-115-308957	Sequence 308957, A	797	37	38.5	326	4	US-10-437-963-127390	Sequence 127390,
725	37	38.5	107	4	US-10-424-599-169314	Sequence 169314,	798	37	38.5	334	4	US-10-425-114-46477	Sequence 46477, A
726	37	38.5	107	5	US-10-723-434-44	Sequence 44, Appl	799	37	38.5	336	4	US-10-437-963-163746	Sequence 163746,
727	37	38.5	108	5	US-10-805-177-40	Sequence 16, Appl	800	37	38.5	348	5	US-10-732-923-6496	Sequence 6496, Ap
728	37	38.5	110	4	US-10-437-963-199662	Sequence 199662,	801	37	38.5	351	4	US-10-437-963-187465	Sequence 187465,
729	37	38.5	111	4	US-10-437-963-104095	Sequence 104095,	802	37	38.5	354	4	US-10-425-114-47467	Sequence 47467, A
730	37	38.5	115	4	US-10-425-115-308957	Sequence 308957, A	803	37	38.5	364	3	US-09-738-626-6869	Sequence 6869, Ap
731	37	38.5	122	5	US-10-450-763-39260	Sequence 39260, A	804	37	38.5	365	4	US-10-344-738-52	Sequence 52, Appl
732	37	38.5	129	4	US-10-424-599-211280	Sequence 211280,	805	37	38.5	368	4	US-10-437-963-202686	Sequence 202686,
733	37	38.5	130	5	US-10-450-763-47628	Sequence 47628, A	806	37	38.5	369	6	US-11-097-143-35580	Sequence 35580, A
734	37	38.5	133	4	US-10-156-761-15084	Sequence 15084, A	807	37	38.5	370	4	US-10-425-114-58999	Sequence 58999, A
735	37	38.5	133	4	US-10-425-115-263718	Sequence 263718, A	808	37	38.5	373	4	US-10-627-476-276	Sequence 276, App
736	37	38.5	134	4	US-10-087-132-1089	Sequence 1089, Ap	809	37	38.5	376	4	US-10-138-714-1058	Sequence 3058, Ap
737	37	38.5	140	4	US-10-282-122A-45368	Sequence 45368, A	810	37	38.5	384	4	US-10-425-114-40373	Sequence 40373, A
738	37	38.5	144	4	US-10-437-963-135866	Sequence 135866,	811	37	38.5	384	4	US-10-425-115-236751	Sequence 236751,
739	37	38.5	150	4	US-10-424-599-264254	Sequence 264254,	812	37	38.5	388	4	US-10-282-122A-50421	Sequence 50421, A
740	37	38.5	156	5	US-10-805-177-133	Sequence 133, App	813	37	38.5	391	5	US-10-469-204-134	Sequence 134, App
741	37	38.5	162	4	US-10-425-114-47461	Sequence 47461, A	814	37	38.5	394	4	US-10-437-963-1233955	Sequence 1233955,
742	37	38.5	162	4	US-10-425-114-59620	Sequence 59620, A	815	37	38.5	401	4	US-10-425-115-191193	Sequence 191193,
743	37	38.5	163	4	US-10-425-114-60931	Sequence 60931, A	816	37	38.5	406	4	US-10-156-761-9773	Sequence 9773, Ap
744	37	38.5	164	4	US-10-425-114-60808	Sequence 60808, A	817	37	38.5	422	4	US-10-425-114-36838	Sequence 36838, A
745	37	38.5	165	4	US-10-425-114-42755	Sequence 42755, A	818	37	38.5	425	5	US-10-492-928A-160	Sequence 160, App
746	37	38.5	166	4	US-10-425-114-60786	Sequence 60786, A	819	37	38.5	436	3	US-10-425-115-16193	Sequence 14, Appl
747	37	38.5	166	4	US-10-425-114-58128	Sequence 58128, A	820	37	38.5	436	3	US-09-730-772-14	Sequence 14, Appl
748	37	38.5	166	4	US-10-425-114-60216	Sequence 60216, A	821	37	38.5	436	3	US-09-735-849-14	Sequence 85, Appl
749	37	38.5	166	4	US-10-425-114-61634	Sequence 61634, A	822	37	38.5	436	3	US-09-574-819-14	Sequence 14, Appl
750	37	38.5	166	4	US-10-425-114-67706	Sequence 67706, A	823	37	38.5	436	3	US-09-930-512-72	Sequence 72, Appl
751	37	38.5	166	4	US-10-425-115-225107	Sequence 225107,	824	37	38.5	436	4	US-10-379-830-14	Sequence 14, Appl
752	37	38.5	167	4	US-10-437-963-127392	Sequence 127392,	825	37	38.5	436	4	US-10-164-279-57	Sequence 57, Appl
753	37	38.5	172	4	US-10-425-115-317769	Sequence 317769,	826	37	38.5	436	5	US-10-851-438-85	Sequence 85, Appl
754	37	38.5	173	5	US-10-755-415-322	Sequence 322, App	827	37	38.5	437	4	US-10-282-122A-66388	Sequence 66388, A
755	37	38.5	174	4	US-10-767-701-55256	Sequence 55256, A	828	37	38.5	442	6	US-10-097-143-15627	Sequence 15627, A
756	37	38.5	175	4	US-10-424-599-192565	Sequence 192565,	829	37	38.5	476	4	US-11-097-143-15627	Sequence 2300, Ap
757	37	38.5	177	4	US-10-425-115-346193	Sequence 346193,	830	37	38.5	476	4	US-10-264-049-2300	Sequence 2300, Ap
			183	4	US-10-437-963-106391	Sequence 106391,				478	4	US-10-425-114-62363	Sequence 62363, A

831	37	38.5	487	4	US-10-369-493-17261	Sequence 17261, A	904	37	38.5	835	5.	US-10-756-149-5729	Sequence 5729, Ap
832	37	38.5	487	5	US-10-732-923-9845	Sequence 9845, Ap	905	37	38.5	851	4	US-10-369-493-2313	Sequence 2313, Ap
833	37	38.5	489	4	US-10-128-714-8058	Sequence 8058, Ap	906	37	38.5	851	5	US-10-732-923-22100	Sequence 22100, A
834	37	38.5	500	4	US-10-437-963-204298	Sequence 204298, A	907	37	38.5	882	4	US-10-298-417-2	Sequence 2, Appli
835	37	38.5	504	4	US-10-437-963-128412	Sequence 128412, A	908	37	38.5	883	4	US-10-322-696-6	Sequence 6, Appli
836	37	38.5	506	4	US-10-425-115-116847	Sequence 316847, A	909	37	38.5	910	5	US-10-466-992-22	Sequence 22, Appli
837	37	38.5	508	4	US-10-686-947-220	Sequence 220, App	910	37	38.5	912	4	US-10-369-493-16719	Sequence 16719, A
838	37	38.5	508	5	US-10-943-507-217	Sequence 217, App	911	37	38.5	920	4	US-10-104-047-2574	Sequence 2574, Ap
839	37	38.5	516	4	US-10-425-115-254858	Sequence 254858, A	912	37	38.5	925	4	US-10-369-493-9234	Sequence 9234, Ap
840	37	38.5	518	4	US-10-437-963-140029	Sequence 140029, A	913	37	38.5	927	4	US-10-369-493-21184	Sequence 21184, A
841	37	38.5	518	4	US-10-686-947-172	Sequence 172, App	914	37	38.5	931	4	US-10-369-493-15855	Sequence 15855, A
842	37	38.5	518	4	US-10-686-947-190	Sequence 190, App	915	37	38.5	931	4	US-10-369-493-16230	Sequence 16230, A
843	37	38.5	518	5	US-10-943-507-170	Sequence 170, App	916	37	38.5	931	4	US-10-282-122A-66953	Sequence 66953, A
844	37	38.5	532	4	US-10-437-963-204310	Sequence 204310, A	917	37	38.5	933	3	US-09-815-242-10095	Sequence 10095, A
845	37	38.5	532	5	US-10-473-127-395	Sequence 395, App	918	37	38.5	933	3	US-09-815-242-14055	Sequence 14055, A
846	37	38.5	539	5	US-10-473-127-396	Sequence 396, App	919	37	38.5	933	4	US-10-078-107-5	Sequence 5, Appli
847	37	38.5	539	5	US-10-473-127-395	Sequence 395, App	920	37	38.5	933	4	US-10-077-751-5	Sequence 5, Appli
848	37	38.5	539	5	US-10-473-127-397	Sequence 397, App	921	37	38.5	933	4	US-10-315-023-9	Sequence 9, Appli
849	37	38.5	539	5	US-10-473-127-398	Sequence 398, App	922	37	38.5	933	4	US-10-077-745-5	Sequence 5, Appli
850	37	38.5	549	4	US-10-032-585-7181	Sequence 7181, Ap	923	37	38.5	933	4	US-10-338-915-5	Sequence 5, Appli
851	37	38.5	562	4	US-10-437-963-135179	Sequence 135179, A	924	37	38.5	933	4	US-10-369-493-758	Sequence 758, App
852	37	38.5	569	6	US-11-097-143-21414	Sequence 21414, A	925	37	38.5	933	4	US-10-282-122A-56482	Sequence 56482, A
853	37	38.5	587	5	US-10-450-763-57378	Sequence 57378, A	926	37	38.5	933	4	US-10-282-122A-75044	Sequence 75044, A
854	37	38.5	593	4	US-10-425-115-298951	Sequence 298951, A	927	37	38.5	933	4	US-10-282-122A-75765	Sequence 75765, A
855	37	38.5	594	5	US-10-631-467-1401	Sequence 1401, Ap	928	37	38.5	934	4	US-10-369-493-397	Sequence 397, App
856	37	38.5	594	5	US-10-631-467-1476	Sequence 1476, Ap	929	37	38.5	935	3	US-09-784-208-3	Sequence 3, Appli
857	37	38.5	596	4	US-10-104-047-2541	Sequence 2541, Ap	930	37	38.5	935	4	US-10-078-107-1	Sequence 1, Appli
858	37	38.5	596	4	US-10-425-114-38308	Sequence 38308, A	931	37	38.5	935	4	US-10-077-751-1	Sequence 1, Appli
859	37	38.5	635	4	US-10-282-122A-55328	Sequence 55328, A	932	37	38.5	935	4	US-10-315-023-3	Sequence 3, Appli
860	37	38.5	646	6	US-11-097-143-7512	Sequence 7512, Ap	933	37	38.5	935	4	US-10-315-023-8	Sequence 8, Appli
861	37	38.5	651	4	US-10-437-963-195137	Sequence 195137, A	934	37	38.5	935	4	US-10-077-745-1	Sequence 1, Appli
862	37	38.5	661	4	US-10-425-115-254542	Sequence 254542, A	935	37	38.5	935	4	US-10-338-915-1	Sequence 1, Appli
863	37	38.5	673	4	US-10-287-274-328	Sequence 328, App	936	37	38.5	935	4	US-10-282-122A-55890	Sequence 55890, A
864	37	38.5	673	4	US-10-369-493-23617	Sequence 23617, A	937	37	38.5	935	4	US-10-282-122A-59453	Sequence 59453, A
865	37	38.5	673	4	US-10-282-122A-42655	Sequence 42655, A	938	37	38.5	935	4	US-10-282-122A-78187	Sequence 78187, A
866	37	38.5	673	4	US-10-282-122A-73457	Sequence 73457, A	939	37	38.5	936	4	US-11-150-265-3	Sequence 3, Appli
867	37	38.5	674	4	US-10-282-122A-74964	Sequence 74964, A	940	37	38.5	936	4	US-10-282-122A-73361	Sequence 73361, A
868	37	38.5	674	4	US-10-282-122A-75806	Sequence 75806, A	941	37	38.5	936	4	US-10-282-122A-77388	Sequence 77388, A
869	37	38.5	675	6	US-11-097-143-13998	Sequence 13998, A	942	37	38.5	936	6	US-11-150-265-8	Sequence 8, Appli
870	37	38.5	682	5	US-10-450-763-40949	Sequence 40949, A	943	37	38.5	949	4	US-10-369-493-17836	Sequence 17836, A
871	37	38.5	686	4	US-10-369-493-22347	Sequence 22347, A	944	37	38.5	950	3	US-09-815-242-11271	Sequence 11271, A
872	37	38.5	699	4	US-10-369-493-4524	Sequence 4524, Ap	945	37	38.5	950	4	US-10-282-122A-58576	Sequence 58576, A
873	37	38.5	705	4	US-10-369-493-7283	Sequence 7283, Ap	946	37	38.5	953	4	US-10-369-493-15484	Sequence 15484, A
874	37	38.5	707	4	US-10-620-669-18	Sequence 18, Appl	947	37	38.5	953	4	US-10-282-122A-46672	Sequence 46672, A
875	37	38.5	746	4	US-10-142-143-19	Sequence 19, Appl	948	37	38.5	956	4	US-10-369-493-10024	Sequence 10024, A
876	37	38.5	746	4	US-10-642-248-65	Sequence 65, Appl	949	37	38.5	957	4	US-10-369-493-17603	Sequence 17603, A
877	37	38.5	758	4	US-10-437-963-187468	Sequence 187468, A	950	37	38.5	1023	5	US-10-486-977-3	Sequence 3, Appli
878	37	38.5	773	4	US-10-280-862-39	Sequence 39, Appl	951	37	38.5	1061	4	US-10-437-963-135175	Sequence 135175, A
879	37	38.5	773	4	US-10-298-680-39	Sequence 39, Appl	952	37	38.5	1063	4	US-10-377-035-15	Sequence 15, Appl
880	37	38.5	773	4	US-10-702-400-49	Sequence 49, Appl	953	37	38.5	1106	4	US-10-282-122A-67531	Sequence 67531, A
881	37	38.5	773	5	US-10-805-650-49	Sequence 49, Appl	954	37	38.5	1134	3	US-09-836-392-16	Sequence 16, Appl
882	37	38.5	774	4	US-10-280-962-41	Sequence 41, Appl	955	37	38.5	1140	6	US-11-097-143-33987	Sequence 33987, A
883	37	38.5	774	4	US-10-298-680-41	Sequence 41, Appl	956	37	38.5	1167	5	US-10-794-514A-328	Sequence 328, App
884	37	38.5	774	4	US-10-702-400-51	Sequence 51, Appl	957	37	38.5	1189	4	US-10-369-493-8120	Sequence 8120, Ap
885	37	38.5	774	5	US-10-805-650-51	Sequence 51, Appl	958	37	38.5	1237	3	US-10-863-906-1	Sequence 78, Appl
886	37	38.5	775	4	US-10-280-962-37	Sequence 37, Appl	959	37	38.5	1237	5	US-10-989-228-78	Sequence 1, Appli
887	37	38.5	775	4	US-10-298-680-37	Sequence 37, Appl	960	37	38.5	1237	5	US-10-863-906-1	Sequence 78, Appl
888	37	38.5	775	4	US-10-702-400-47	Sequence 47, Appl	961	37	38.5	1251	4	US-10-437-963-156608	Sequence 156608, A
889	37	38.5	775	5	US-10-805-650-47	Sequence 47, Appl	962	37	38.5	1286	5	US-10-486-977-28	Sequence 28, Appli
890	37	38.5	786	5	US-10-501-282-4436	Sequence 4436, Ap	963	37	38.5	1336	4	US-10-437-963-173565	Sequence 173565, A
891	37	38.5	790	4	US-10-408-765A-2178	Sequence 2178, Ap	964	37	38.5	1528	6	US-11-097-143-2013	Sequence 2013, Ap
892	37	38.5	795	4	US-10-437-963-122057	Sequence 122057, A	965	37	38.5	1529	5	US-10-732-923-8686	Sequence 8686, Ap
893	37	38.5	807	4	US-10-369-493-10414	Sequence 10414, A	966	37	38.5	1547	4	US-10-473-670-12	Sequence 12, Appl
894	37	38.5	820	3	US-09-764-875-633	Sequence 633, App	967	37	38.5	1570	4	US-10-369-022-38	Sequence 38, Appl
895	37	38.5	823	4	US-10-101-464A-923	Sequence 923, App	968	37	38.5	1570	4	US-10-757-262-102	Sequence 102, App
896	37	38.5	823	5	US-10-864-252-923	Sequence 923, App	969	37	38.5	1570	5	US-10-840-512-118	Sequence 118, App
897	37	38.5	825	4	US-10-101-464A-922	Sequence 922, App	970	37	38.5	1583	4	US-10-042-865-78	Sequence 78, Appl
898	37	38.5	825	5	US-10-864-252-922	Sequence 922, App	971	37	38.5	1607	4	US-10-425-115-272588	Sequence 272588, A
899	37	38.5	829	4	US-10-282-122A-77635	Sequence 77635, A	972	37	38.5	1633	4	US-10-029-386-33090	Sequence 33090, A
900	37	38.5	831	3	US-09-789-561-86	Sequence 86, Appl	973	37	38.5	1826	4	US-10-408-765A-1477	Sequence 1477, Ap
901	37	38.5	831	3	US-09-833-245-2153	Sequence 2153, Ap	974	37	38.5	1854	4	US-10-369-493-2233	Sequence 2233, Ap
902	37	38.5	831	5	US-10-883-936-86	Sequence 86, Appl	975	37	38.5	2073	4	US-10-741-191-15	Sequence 15, Appl
903	37	38.5	835	5	US-10-723-860-4161	Sequence 4161, Ap	976	37	38.5	2073	4	US-10-742-350-15	Sequence 15, Appl

```
977 37 38.5 2081 4 US-10-369-493-3423 Sequence 3423, Ap
978 37 38.5 2199 4 US-10-437-963-168662 Sequence 168662,
979 37 38.5 2303 4 US-10-267-502-371 Sequence 371, App
980 37 38.5 2303 5 US-10-852-335A-109 Sequence 109, App
981 37 38.5 8601 5 US-10-496-377-7 Sequence 7, Appli
982 37 38.5 9548 5 US-10-819-386A-7 Sequence 7, Appli
983 36.5 38.0 44 4 US-10-450-763-58956 Sequence 58956, A
984 36.5 38.0 85 4 US-10-425-115-262752 Sequence 262752, A
985 36.5 38.0 95 5 US-10-450-763-39915 Sequence 39915, A
986 36.5 38.0 95 5 US-10-450-763-45536 Sequence 45536, A
987 36.5 38.0 95 5 US-10-450-763-57478 Sequence 57478, A
988 36.5 38.0 107 4 US-10-425-115-277553 Sequence 277553,
989 36.5 38.0 107 4 US-10-425-115-319846 Sequence 319846,
990 36.5 38.0 109 4 US-10-424-599-203285 Sequence 203285,
991 36.5 38.0 110 4 US-10-767-701-60482 Sequence 60482, A
992 36.5 38.0 149 4 US-10-767-701-57894 Sequence 57894, A
993 36.5 38.0 149 5 US-10-450-763-44497 Sequence 44497, A
994 36.5 38.0 197 4 US-10-781-014-526 Sequence 526, App
995 36.5 38.0 222 5 US-10-450-763-36890 Sequence 36890, A
996 36.5 38.0 322 5 US-11-097-143-25536 Sequence 25536, A
997 36.5 38.0 335 3 US-09-738-626-4330 Sequence 4330, Ap
998 36.5 38.0 370 6 US-11-097-143-11082 Sequence 11082, A
999 36.5 38.0 418 5 US-10-732-923-18868 Sequence 18868, A
1000 36.5 38.0 432 4 US-10-119-651-8 Sequence 8, Appli
```

## ALIGNMENTS

```
RESULT 1
US-10-390-566-4
; Sequence 4, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; PRIOR FILING DATE: 2003-03-17
; PRIOR FILING DATE: 2003-03-17
; PRIOR FILING DATE: 2003-03-17
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-4

Query Match 100.0%; Score 96; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 9e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSSIWPQEEPAIRPR 17
Db 6 FRDSSIWPQEEPAIRPR 22

RESULT 2
US-10-390-566-11
; Sequence 11, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
```

```
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; PRIOR FILING DATE: 2003-03-17
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-11

Query Match 100.0%; Score 96; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 9e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSSIWPQEEPAIRPR 17
Db 6 FRDSSIWPQEEPAIRPR 22

RESULT 3
US-10-390-566-17
; Sequence 17, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; PRIOR FILING DATE: 2003-03-17
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-17

Query Match 100.0%; Score 96; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 9e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSSIWPQEEPAIRPR 17
Db 6 FRDSSIWPQEEPAIRPR 22

RESULT 4
US-10-390-566-24
; Sequence 24, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
```

```
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-24

Query Match      100.0%; Score 96; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 9e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSSIWPQEEPAIRPR 17
   |||||
Db 6 FRDSSIWPQEEPAIRPR 22

RESULT 5
US-10-390-566-3
; Sequence 3, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-3

Query Match      100.0%; Score 96; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 9.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSSIWPQEEPAIRPR 17
   |||||
Db 16 FRDSSIWPQEEPAIRPR 32

RESULT 6
US-10-390-566-10
; Sequence 10, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-10
```

```
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-10

Query Match      100.0%; Score 96; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 9.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSSIWPQEEPAIRPR 17
   |||||
Db 16 FRDSSIWPQEEPAIRPR 32

RESULT 7
US-10-390-566-16
; Sequence 16, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-16

Query Match      100.0%; Score 96; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 9.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSSIWPQEEPAIRPR 17
   |||||
Db 16 FRDSSIWPQEEPAIRPR 32

RESULT 8
US-10-390-566-23
; Sequence 23, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-23
```



```
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-23

Query Match      100.0%; Score 96; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 9.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSSIWPQEEPAIRPR 17
   |||||
Db 16 FRDSSIWPQEEPAIRPR 32
   |||||

RESULT 9
US-10-665-602-7
; Sequence 7, Application US/10665602
; Publication No. US20040086967A1
; GENERAL INFORMATION:
; APPLICANT: Meissner, Paul S.
; Coleman, Timothy A.
; TITLE OF INVENTION: Human Cryptin Growth Factor
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/665,602
; FILING DATE: 22-Sep-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/393,023A
; FILING DATE: 09-SEP-1999
; APPLICATION NUMBER: US 08/471,371
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Marks, Michelle S.
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PF200D1
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-665-602-7

Query Match      100.0%; Score 96; DB 4; Length 174;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSSIWPQEEPAIRPR 17
   |||||
Db 32 FRDSSIWPQEEPAIRPR 48
   |||||
```

```
RESULT 10
US-10-225-486-66
; Sequence 66, Application US/10225486
; Publication No. US20030129207A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Secrist, Heather
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.591
; CURRENT APPLICATION NUMBER: US/10/225,486
; CURRENT FILING DATE: 2002-10-29
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 66
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-486-66

Query Match      100.0%; Score 96; DB 4; Length 187;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSSIWPQEEPAIRPR 17
   |||||
Db 46 FRDSSIWPQEEPAIRPR 62
   |||||

RESULT 11
US-10-241-220-58
; Sequence 58, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Prantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 58
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-241-220-58

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSSIWPQEEPAIRPR 17
   |||||
Db 46 FRDSSIWPQEEPAIRPR 62
   |||||

RESULT 12
US-10-388-578-6
; Sequence 6, Application US/10388578
; Publication No. US2003024411A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Stanton, Lawrence
```



```
; APPLICANT: Ralph, Brandenberger
; APPLICANT: Joseph, Gold D.
; APPLICANT: John, Irving
; APPLICANT: Mandalam, Ramkumar
; APPLICANT: Mok, Michael
; APPLICANT: Sheiton, Dawne
; TITLE OF INVENTION: Genes that are Up- or Down-Regulated During Differentiation of Hu
; FILE REFERENCE: 135/001
; CURRENT APPLICATION NUMBER: US/10/388,578
; PRIOR FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: Custom
; SEQ ID NO 6
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-388-578-6

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSIWPEEPAIRPR 17
Db      46 FRDSIWPEEPAIRPR 62

RESULT 13
US-10-390-566-1
; Sequence 1, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; PRIOR FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-1

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSIWPEEPAIRPR 17
Db      46 FRDSIWPEEPAIRPR 62

RESULT 14
US-10-390-566-2
; Sequence 2, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; PRIOR FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-9

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSIWPEEPAIRPR 17
Db      46 FRDSIWPEEPAIRPR 62

RESULT 15
US-10-390-566-9
; Sequence 9, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; PRIOR FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-9

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSIWPEEPAIRPR 17
Db      46 FRDSIWPEEPAIRPR 62

RESULT 16
US-10-390-566-22
; Sequence 22, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
```

```
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-2

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSIWPEEPAIRPR 17
Db      46 FRDSIWPEEPAIRPR 62

RESULT 15
US-10-390-566-9
; Sequence 9, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; PRIOR FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-9

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSIWPEEPAIRPR 17
Db      46 FRDSIWPEEPAIRPR 62

RESULT 16
US-10-390-566-22
; Sequence 22, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
```

```
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-22

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSDIWPQEEPAIRPR 17
Db      46 FRDSDIWPQEEPAIRPR 62

RESULT 17
US-10-362-597A-3
; Sequence 3, Application US/10362597A
; Publication No. US20040054142A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237
; CURRENT APPLICATION NUMBER: US/10/362,597A
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: PCT/EP01/0964
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Human
US-10-362-597A-3

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSDIWPQEEPAIRPR 17
Db      46 FRDSDIWPQEEPAIRPR 62

RESULT 18
US-10-362-597A-4
; Sequence 4, Application US/10362597A
; Publication No. US20040054142A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237
; CURRENT APPLICATION NUMBER: US/10/362,597A
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: PCT/EP01/0964
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 188
```

```
; TYPE: PRT
; ORGANISM: human
US-10-362-597A-4

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSDIWPQEEPAIRPR 17
Db      46 FRDSDIWPQEEPAIRPR 62

RESULT 19
US-10-362-597A-6
; Sequence 6, Application US/10362597A
; Publication No. US20040054142A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237
; CURRENT APPLICATION NUMBER: US/10/362,597A
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: PCT/EP01/0964
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 188
; TYPE: PRT
; ORGANISM: human
US-10-362-597A-6

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSDIWPQEEPAIRPR 17
Db      46 FRDSDIWPQEEPAIRPR 62

RESULT 20
US-10-362-597A-96
; Sequence 96, Application US/10362597A
; Publication No. US20040054142A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237
; CURRENT APPLICATION NUMBER: US/10/362,597A
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: PCT/EP01/0964
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 188
; TYPE: PRT
; ORGANISM: human
US-10-362-597A-96

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSDIWPQEEPAIRPR 17
Db      46 FRDSDIWPQEEPAIRPR 62
```

```
Db      |||||||
46 FRDSIWQEEPAIRPR 62

RESULT 21
US-10-407-481-3
; Sequence 3, Application US/10407481
; Publication No. US20040138112A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237-1
; CURRENT APPLICATION NUMBER: US/10/407,481
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: PCT/EP01/09646
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Human
US-10-407-481-3

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSIWQEEPAIRPR 17
Db      |||||||
46 FRDSIWQEEPAIRPR 62

RESULT 22
US-10-407-481-4
; Sequence 4, Application US/10407481
; Publication No. US20040138112A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237-1
; CURRENT APPLICATION NUMBER: US/10/407,481
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: PCT/EP01/09646
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 188
; TYPE: PRT
; ORGANISM: human
US-10-407-481-4

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSIWQEEPAIRPR 17
Db      |||||||
46 FRDSIWQEEPAIRPR 62

RESULT 23
US-10-407-481-6
; Sequence 6, Application US/10407481
; Publication No. US20040138112A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237-1
; CURRENT APPLICATION NUMBER: US/10/407,481
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: PCT/EP01/09646
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 188
; TYPE: PRT
; ORGANISM: human
US-10-407-481-6

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSIWQEEPAIRPR 17
Db      |||||||
46 FRDSIWQEEPAIRPR 62

RESULT 24
US-10-407-481-96
; Sequence 96, Application US/10407481
; Publication No. US20040138112A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237-1
; CURRENT APPLICATION NUMBER: US/10/407,481
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: PCT/EP01/09646
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 188
; TYPE: PRT
; ORGANISM: human
US-10-407-481-96

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSIWQEEPAIRPR 17
Db      |||||||
46 FRDSIWQEEPAIRPR 62

RESULT 25
US-10-712-124-58
; Sequence 58, Application US/10712124
; Publication No. US20040146907A1
; GENERAL INFORMATION:
; APPLICANT: SMITH, VICTORIA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING DYSPLASIA
; FILE REFERENCE: P2000R1
; CURRENT APPLICATION NUMBER: US/10/712,124
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 60/425,813
; PRIOR FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 123
; SEQ ID NO 58
```

```
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-712-124-58

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 26
US-10-693-538-1
; Sequence 1, Application US/10693538
; Publication No. US20040146940A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Sanicola-Nadel, Michele
; APPLICANT: Williams, Kevin
; APPLICANT: Schiffer, Susan
; APPLICANT: Rayhorn, Paul
; TITLE OF INVENTION: Cripoto Blocking Antibodies and Uses
; FILE REFERENCE: BIN-A117CN
; CURRENT APPLICATION NUMBER: US/10/693,538
; CURRENT FILING DATE: 2003-10-23
; PRIOR APPLICATION NUMBER: PCT/US02/11950
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: 60/286,782
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/293,020
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/301,091
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/367,002
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-693-538-1

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 27
US-10-693-538-2
; Sequence 2, Application US/10693538
; Publication No. US20040146940A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Sanicola-Nadel, Michele
; APPLICANT: Williams, Kevin
; APPLICANT: Schiffer, Susan
; APPLICANT: Rayhorn, Paul
; TITLE OF INVENTION: Cripoto Blocking Antibodies and Uses
; FILE REFERENCE: BIN-A117CN
; CURRENT APPLICATION NUMBER: US/10/693,538
; CURRENT FILING DATE: 2003-10-23
; PRIOR APPLICATION NUMBER: PCT/US02/11950
; PRIOR FILING DATE: 2002-04-16
```

```
; PRIOR APPLICATION NUMBER: 60/286,782
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/293,020
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/301,091
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/367,002
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-693-538-2

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 28
US-10-389-431-6
; Sequence 6, Application US/10389431
; Publication No. US20040180347A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Stanton, Lawrence
; APPLICANT: Ralph, Brandenberger
; APPLICANT: Joseph, Gold D.
; APPLICANT: John, Irving
; APPLICANT: Mandalam, Ramkumar
; APPLICANT: Mok, Michael
; TITLE OF INVENTION: A Marker System for Preparing and Characterizing High-Quality Human Embryonic Stem Cells
; FILE REFERENCE: 135/002
; CURRENT APPLICATION NUMBER: US/10/389,431
; CURRENT FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-389-431-6

Query Match      100.0%; Score 96; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 29
US-10-816-476-3
; Sequence 3, Application US/10816476
; Publication No. US20040202648A1
; GENERAL INFORMATION:
; APPLICANT: CABEZON, Teresa Elisa Virginia Silva
; APPLICANT: GERARD, Catherine Marie Ghislaine
; APPLICANT: PALMANTIER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: VU60111
; CURRENT APPLICATION NUMBER: US/10/816,476
; CURRENT FILING DATE: 2004-04-01
; PRIOR APPLICATION NUMBER:
```

; PRIOR FILING DATE: 2004-04-01  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Human  
US-10-816-476-3

Query Match 100.0%; Score 96; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSSIWPQEEPAIRPR 17  
Db 46 FRDSSIWPQEEPAIRPR 62

## RESULT 30

US-10-816-476-4  
; Sequence 4, Application US/10816476  
; Publication No. US20040202648A1  
; GENERAL INFORMATION:  
; APPLICANT: CABEZON, Teresa Elisa Virginia Silva  
; APPLICANT: GERARD, Catherine Marie Ghislaine  
; APPLICANT: PALMANTIER, Remi M.  
; APPLICANT: VINALS Y DE BASSOLS, Carlota  
; TITLE OF INVENTION: Vaccines  
; FILE REFERENCE: VU60111  
; CURRENT APPLICATION NUMBER: US/10/816,476  
; PRIOR FILING DATE: 2004-04-01  
; PRIOR APPLICATION NUMBER:  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: human  
US-10-816-476-4

Query Match 100.0%; Score 96; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSSIWPQEEPAIRPR 17  
Db 46 FRDSSIWPQEEPAIRPR 62

## RESULT 31

US-10-816-476-6  
; Sequence 6, Application US/10816476  
; Publication No. US20040202648A1  
; GENERAL INFORMATION:  
; APPLICANT: CABEZON, Teresa Elisa Virginia Silva  
; APPLICANT: GERARD, Catherine Marie Ghislaine  
; APPLICANT: PALMANTIER, Remi M.  
; APPLICANT: VINALS Y DE BASSOLS, Carlota  
; TITLE OF INVENTION: Vaccines  
; FILE REFERENCE: VU60111  
; CURRENT APPLICATION NUMBER: US/10/816,476  
; PRIOR FILING DATE: 2004-04-01  
; PRIOR APPLICATION NUMBER:  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: human  
US-10-816-476-6

Query Match 100.0%; Score 96; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSSIWPQEEPAIRPR 17  
Db 46 FRDSSIWPQEEPAIRPR 62

## RESULT 32

US-10-816-476-96  
; Sequence 96, Application US/10816476  
; Publication No. US20040202648A1  
; GENERAL INFORMATION:  
; APPLICANT: CABEZON, Teresa Elisa Virginia Silva  
; APPLICANT: GERARD, Catherine Marie Ghislaine  
; APPLICANT: PALMANTIER, Remi M.  
; APPLICANT: VINALS Y DE BASSOLS, Carlota  
; TITLE OF INVENTION: Vaccines  
; FILE REFERENCE: VU60111  
; CURRENT APPLICATION NUMBER: US/10/816,476  
; PRIOR FILING DATE: 2004-04-01  
; PRIOR APPLICATION NUMBER:  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 96  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: human  
US-10-816-476-96

Query Match 100.0%; Score 96; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSSIWPQEEPAIRPR 17  
Db 46 FRDSSIWPQEEPAIRPR 62

## RESULT 33

US-10-872-972-58  
; Sequence 58, Application US/10872972  
; Publication No. US20040229277A1  
; GENERAL INFORMATION:  
; APPLICANT: Prantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P5010R1-US  
; CURRENT APPLICATION NUMBER: US/10/872,972  
; CURRENT FILING DATE: 2004-06-21  
; PRIOR FILING DATE: 2004-06-21  
; PRIOR APPLICATION NUMBER: US/10/241,220  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 58  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-872-972-58

Query Match 100.0%; Score 96; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSSIWPQEEPAIRPR 17

```
Db 46 FRDSIWQEEPAIRPR 62
|||||
RESULT 34
US-10-872-991-58
; Sequence 58, Application US/10872991
; Publication No. US20040242860A1
; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/872,991
; CURRENT FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US/10/241,220
; PRIOR FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 58
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-872-991-58

Query Match 100.0%; Score 96; DB 5; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWQEEPAIRPR 17
|||||
Db 46 FRDSIWQEEPAIRPR 62

RESULT 35
US-10-491-997-28
; Sequence 28, Application US/10491997
; Publication No. US20050089957A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TREATMENT OF INFLAMMATORY BOWEL DISORDERS
; FILE REFERENCE: P1915R1 US
; CURRENT APPLICATION NUMBER: US/10/491,997
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: PCT/US02/33070
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/340,083
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 28
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-491-997-28

Query Match 100.0%; Score 96; DB 5; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWQEEPAIRPR 17
|||||
Db 46 FRDSIWQEEPAIRPR 62

RESULT 36
US-10-983-340-13
; Sequence 13, Application US/10983340
; Publication No. US20050238649A1
; GENERAL INFORMATION:
; APPLICANT: Doronina, Svetlana O.
; APPLICANT: Toki, Brian E.
; APPLICANT: Senter, Peter D.
```

```
US-10-491-997-72
; Sequence 72, Application US/10491997
; Publication No. US20050089957A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TREATMENT OF INFLAMMATORY BOWEL DISORDERS
; FILE REFERENCE: P1915R1 US
; CURRENT APPLICATION NUMBER: US/10/491,997
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: PCT/US02/33070
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/340,083
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 72
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-491-997-72

Query Match 100.0%; Score 96; DB 5; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWQEEPAIRPR 17
|||||
Db 46 FRDSIWQEEPAIRPR 62

RESULT 37
US-10-940-431-2
; Sequence 2, Application US/10940431
; Publication No. US20050208045A1
; GENERAL INFORMATION:
; APPLICANT: Vale, Wylie
; APPLICANT: Harrison, Craig A.
; APPLICANT: Gray, Peter C.
; TITLE OF INVENTION: Cryptot Antagonism of Activin and TGF- $\beta$ 
; FILE REFERENCE: D6525
; CURRENT APPLICATION NUMBER: US/10/940,431
; CURRENT FILING DATE: 2004-09-14
; PRIOR APPLICATION NUMBER: 60/503,046
; PRIOR FILING DATE: 2003-09-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Macintosh OS 10
; SEQ ID NO 2
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: amino acid sequence of human Crypto protein
US-10-940-431-2

Query Match 100.0%; Score 96; DB 5; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWQEEPAIRPR 17
|||||
Db 46 FRDSIWQEEPAIRPR 62

RESULT 38
US-10-983-340-13
; Sequence 13, Application US/10983340
; Publication No. US20050238649A1
; GENERAL INFORMATION:
; APPLICANT: Doronina, Svetlana O.
; APPLICANT: Toki, Brian E.
; APPLICANT: Senter, Peter D.
```

; APPLICANT: Ebens, Allen J.  
; APPLICANT: Polakis, Paul  
; APPLICANT: Sliwowski, Mark X.  
; APPLICANT: Spencer, Susan D.  
; APPLICANT: Kline, Toni Beth  
; TITLE OF INVENTION: MONOMETHYLVALINE COMPOUNDS CAPABLE OF CONJUGATION TO LIGANDS  
; FILE REFERENCE: 018891-001020US  
; CURRENT APPLICATION NUMBER: US/10/983,340  
; PRIOR FILING DATE: 2004-11-05  
; PRIOR APPLICATION NUMBER: US 60/598,899  
; PRIOR FILING DATE: 2004-08-04  
; PRIOR APPLICATION NUMBER: US 60/557,116  
; PRIOR FILING DATE: 2004-03-26  
; PRIOR APPLICATION NUMBER: US 60/518,534  
; PRIOR FILING DATE: 2003-11-06  
; NUMBER OF SEQ ID NOS: 35  
; SEQ ID NO 13  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-983-340-13

Query Match 100.0%; Score 96; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWPQEEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 46 FRDSIWPQEEPAIRPR 62

RESULT 39  
US-10-390-566-7  
; Sequence 7, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: CRYPTO MUTANT AND USES THEREOF  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; PRIOR FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0  
; SEQ ID NO 7  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-390-566-7

Query Match 100.0%; Score 96; DB 4; Length 360;  
Best Local Similarity 100.0%; Pred. No. 2.4e-05; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWPQEEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 9 FRDSIWPQEEPAIRPR 25

RESULT 40  
US-10-390-566-14  
; Sequence 14, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan

; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: CRYPTO MUTANT AND USES THEREOF  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; PRIOR FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0  
; SEQ ID NO 14  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-390-566-14

Query Match 100.0%; Score 96; DB 4; Length 360;  
Best Local Similarity 100.0%; Pred. No. 2.4e-05; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWPQEEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 9 FRDSIWPQEEPAIRPR 25

RESULT 41  
US-10-390-566-20  
; Sequence 20, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: CRYPTO MUTANT AND USES THEREOF  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; PRIOR FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0  
; SEQ ID NO 20  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-390-566-20

Query Match 100.0%; Score 96; DB 4; Length 360;  
Best Local Similarity 100.0%; Pred. No. 2.4e-05; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWPQEEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 9 FRDSIWPQEEPAIRPR 25

RESULT 42  
US-10-390-566-27  
; Sequence 27, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele

; TITLE OF INVENTION: Cripto Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-27

Query Match 100.0%; Score 96; DB 4; Length 360;  
Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWQEEPAIRPR 17  
Db 9 FRDSDIWQEEPAIRPR 25

## RESULT 43

US-10-390-566-6  
; Sequence 6, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: Cripto Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 367  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-6

Query Match 100.0%; Score 96; DB 4; Length 367;  
Best Local Similarity 100.0%; Pred. No. 2.5e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWQEEPAIRPR 17  
Db 16 FRDSDIWQEEPAIRPR 32

## RESULT 44

US-10-390-566-13  
; Sequence 13, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: Cripto Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566

; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 367  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-13

Query Match 100.0%; Score 96; DB 4; Length 367;  
Best Local Similarity 100.0%; Pred. No. 2.5e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWQEEPAIRPR 17  
Db 16 FRDSDIWQEEPAIRPR 32

## RESULT 45

US-10-390-566-19  
; Sequence 19, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: Cripto Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US01/29066  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19  
; LENGTH: 367  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-390-566-19

Query Match 100.0%; Score 96; DB 4; Length 367;  
Best Local Similarity 100.0%; Pred. No. 2.5e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWQEEPAIRPR 17  
Db 16 FRDSDIWQEEPAIRPR 32

## RESULT 46

US-10-390-566-26  
; Sequence 26, Application US/10390566  
; Publication No. US20030232755A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Kevin P.  
; APPLICANT: Foley, Susan  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Domon, Bruno  
; APPLICANT: Sanicola-Nadel, Michele  
; TITLE OF INVENTION: Cripto Mutant and Uses Thereof  
; FILE REFERENCE: BGN-A104  
; CURRENT APPLICATION NUMBER: US/10/390,566  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/233,148  
; PRIOR FILING DATE: 2000-09-18



```

; NAME/KEY: unsure
; LOCATION: (1)..(166)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51223C.1.pep
US-10-424-599-231675

Query Match      66.7%; Score 64; DB 4; Length 166;
Best Local Similarity 68.8%; Pred. No. 0.36;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      2 RDSIWPQEPFAIRPR 17
         ||:|||| |:||||:|
DB      31 RDNSIWDQKEPAVRDR 46

RESULT 49
US-10-940-431-1
; Sequence 1, Application US/10940431
; Publication No. US20050208045A1
; GENERAL INFORMATION:
; APPLICANT: Vale, Wylie
; APPLICANT: Harrison, Craig A.
; APPLICANT: Gray, Peter C.
; TITLE OF INVENTION: Crypto Antagonism of Activin and TGF- $\beta$ 
; TITLE OF INVENTION: Signaling
; FILE REFERENCE: D6525
; CURRENT APPLICATION NUMBER: US/10/940,431
; CURRENT FILING DATE: 2004-09-14
; PRIOR APPLICATION NUMBER: 60/503,046
; PRIOR FILING DATE: 2003-09-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Macintosh OS 10
; SEQ ID NO 1
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: amino acid sequence of mouse Crypto protein
US-10-940-431-1

Query Match      66.7%; Score 64; DB 5; Length 171;
Best Local Similarity 68.8%; Pred. No. 0.37;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      2 RDSIWPQEPFAIRPR 17
         ||:|||| |:||||:|
DB      31 RDNSIWDQKEPAVRDR 46

RESULT 50
US-10-362-597A-80
; Sequence 80, Application US/10362597A
; Publication No. US20040054142A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237
; CURRENT APPLICATION NUMBER: US/10/362,597A
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: PCT/EP01/0964
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 9
; TYPE: PRT
; ORGANISM: human
US-10-362-597A-80

```

Query Match 56.2%; Score 54; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDDSIWPQ 9  
| | | | | | | | |  
Db 1 FRDDSIWPQ 9

Search completed: February 3, 2006, 19:02:06  
Job time : 132.5 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:56:32 ; Search time 34.5 Seconds

(without alignments)

40.739 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_46\_62

Perfect score: 96

Sequence: 1 FRDSTWQEPRIAPR 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/iaa/5\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/6\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/H\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/PCITUS\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/RE\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	68	1	US-08-208-008C-13
2	96	100.0	174	6	5256643-3
3	96	100.0	187	6	5256643-2
4	96	100.0	188	1	US-07-749-001-3
5	96	100.0	188	1	US-07-749-001-5
6	96	100.0	188	1	US-08-154-198-3
7	96	100.0	188	1	US-08-154-198-5
8	96	100.0	188	1	US-08-463-335-3
9	96	100.0	188	1	US-08-463-335-5
10	96	100.0	188	1	US-08-464-023A-3
11	96	100.0	188	1	US-08-464-023A-5
12	96	100.0	188	1	US-08-471-371-7
13	96	100.0	192	2	US-09-949-016-7062
14	96	100.0	192	2	US-09-949-016-8959
15	96	100.0	192	2	US-09-949-016-11113
16	64	66.7	160	1	US-08-441-629-7
17	64	66.7	160	2	US-08-776-207-7
18	64	66.7	160	2	US-09-507-773-7
19	64	66.7	160	2	US-10-016-447-7
20	64	66.7	160	4	PCT-US95-09172-7
21	46	47.9	252	2	US-09-252-991A-18373
22	45	46.9	386	2	US-09-270-767-43850
23	45	46.9	448	2	US-09-252-991A-30230
24	45	46.9	766	2	US-09-538-092-216
25	44	45.8	240	2	US-09-248-796A-19171
26	43	44.8	1113	2	US-09-894-998A-51
27	43	44.8	1113	2	US-10-237-551-51

28	43	44.8	1114	2	US-10-237-551-202	Sequence 202, App
29	42	43.8	255	2	US-09-489-039A-10445	Sequence 10445, A
30	42	43.8	349	2	US-09-252-991A-29455	Sequence 29455, A
31	42	43.8	779	1	US-08-375-134-12	Sequence 12, Appl
32	42	43.8	779	4	PCT-US95-15263-12	Sequence 12, Appl
33	42	43.8	1736	2	US-09-919-497-98	Sequence 98, Appl
34	42	43.8	1771	2	US-09-949-016-9470	Sequence 9470, Ap
35	41	42.7	226	2	US-09-248-796A-17216	Sequence 17216, A
36	41	42.7	590	1	US-09-489-039A-12982	Sequence 12982, A
37	41	42.7	694	1	US-08-895-522-3	Sequence 3, Appli
38	41	42.7	694	2	US-09-195-391-3	Sequence 3, Appli
39	40	41.7	107	2	US-09-240-374-29	Sequence 29, Appl
40	40	41.7	107	2	US-09-848-798-29	Sequence 29, Appl
41	40	41.7	108	2	US-09-726-219A-244	Sequence 244, App
42	40	41.7	108	2	US-09-196-522-244	Sequence 244, App
43	40	41.7	126	2	US-09-107-532A-5453	Sequence 5453, Ap
44	40	41.7	154	2	US-09-489-039A-8476	Sequence 8476, Ap
45	40	41.7	180	2	US-09-640-211A-834	Sequence 834, App
46	40	41.7	186	2	US-09-252-991A-16962	Sequence 16962, A
47	40	41.7	263	2	US-10-104-047-2017	Sequence 2017, Ap
48	40	41.7	418	2	US-09-252-991A-20824	Sequence 20824, A
49	40	41.7	471	2	US-09-902-540-14431	Sequence 14431, A
50	40	41.7	614	2	US-09-902-540-14372	Sequence 14372, A
51	40	41.7	747	1	US-08-895-522-1	Sequence 1, Appli
52	40	41.7	747	2	US-09-195-391-1	Sequence 1, Appli
53	40	41.7	747	2	US-09-422-840B-2	Sequence 2, Appli
54	40	41.7	747	2	US-09-422-840B-4	Sequence 4, Appli
55	39	40.6	122	2	US-09-270-767-35878	Sequence 35878, A
56	39	40.6	122	2	US-09-270-767-51095	Sequence 51095, A
57	39	40.6	138	2	US-09-252-991A-23683	Sequence 23683, A
58	39	40.6	306	2	US-09-252-991A-30319	Sequence 30319, A
59	39	40.6	316	2	US-09-252-991A-22518	Sequence 22518, A
60	39	40.6	357	2	US-09-724-623-93	Sequence 93, Appl
61	39	40.6	358	2	US-09-949-016-10216	Sequence 10216, A
62	39	40.6	400	2	US-09-902-540-16279	Sequence 16279, A
63	39	40.6	412	2	US-09-252-991A-22952	Sequence 22952, A
64	39	40.6	455	2	US-10-103-547A-2	Sequence 2, Appli
65	39	40.6	466	2	US-09-252-991A-17965	Sequence 17965, A
66	39	40.6	546	2	US-09-902-540-16084	Sequence 16084, A
67	39	40.6	553	1	US-07-683-957B-3	Sequence 3, Appli
68	39	40.6	659	2	US-09-955-732A-2	Sequence 2, Appli
69	39	40.6	687	2	US-09-489-039A-13937	Sequence 13937, A
70	39	40.6	774	2	US-09-803-165-34	Sequence 34, Appl
71	39	40.6	931	2	US-09-949-016-9850	Sequence 9850, Ap
72	39	40.6	1265	2	US-09-862-027-80	Sequence 80, Appl
73	39	40.6	1734	2	US-09-862-027-81	Sequence 81, Appl
74	38.5	40.1	504	2	US-09-519-878-4	Sequence 4, Appli
75	38.5	40.1	615	2	US-09-252-991A-23203	Sequence 23203, A
76	38.5	40.1	1485	2	US-08-851-567B-32	Sequence 32, Appl
77	38.5	40.1	1667	2	US-09-489-039A-11624	Sequence 11624, A
78	38	39.6	39	2	US-09-220-528-48	Sequence 48, Appl
79	38	39.6	107	2	US-09-220-528-52	Sequence 52, Appl
80	38	39.6	108	2	US-08-591-632-68	Sequence 68, Appl
81	38	39.6	108	2	US-09-611-451-68	Sequence 68, Appl
82	38	39.6	132	2	US-09-513-999C-4373	Sequence 4373, Ap
83	38	39.6	163	2	US-09-270-767-35534	Sequence 35534, A
84	38	39.6	163	2	US-09-270-767-50751	Sequence 50751, A
85	38	39.6	207	2	US-09-270-767-50751	Sequence 50751, A
86	38	39.6	220	2	US-09-252-991A-21774	Sequence 21774, A
87	38	39.6	220	2	US-09-220-528-26	Sequence 26, Appl
88	38	39.6	220	2	US-09-347-613C-9	Sequence 9, Appli
89	38	39.6	220	2	US-09-347-613C-35	Sequence 35, Appl
90	38	39.6	220	2	US-09-662-183A-9	Sequence 9, Appli
91	38	39.6	220	2	US-09-662-183A-35	Sequence 35, Appl
92	38	39.6	225	2	US-09-248-335-40	Sequence 40, Appl
93	38	39.6	226	2	US-09-252-991A-29689	Sequence 29689, A
94	38	39.6	231	2	US-09-252-991A-29962	Sequence 29962, A
95	38	39.6	240	2	US-09-489-039A-11311	Sequence 11311, A
96	38	39.6	245	2	US-09-902-540-15908	Sequence 15908, A
97	38	39.6	262	1	US-08-015-973-8	Sequence 8, Appli
98	38	39.6	262	1	US-08-448-164-8	Sequence 8, Appli
99	38	39.6	262	2	US-08-081-529-8	Sequence 8, Appli
100	38	39.6	279	2	US-10-104-047-3382	Sequence 3382, Ap

101	38	39.6	300	2	US-09-252-991A-27295	Sequence 27295, A	174	37.5	39.1	886	2	US-09-784-316-4	Sequence 4, Appli
102	38	39.6	329	2	US-09-413-231-5	Sequence 5, Appli	175	37.5	39.1	885	2	US-10-229-124-4	Sequence 4, Appli
103	38	39.6	329	2	US-09-413-231-5	Sequence 9, Appli	176	37.5	39.1	915	2	US-09-949-016-9697	Sequence 9697, Ap
104	38	39.6	332	2	US-09-252-991A-22640	Sequence 22640, A	177	37	38.5	75	2	US-09-270-767-36728	Sequence 36728, A
105	38	39.6	343	2	US-09-454-034-8	Sequence 8, Appli	178	37	38.5	75	2	US-09-270-767-51945	Sequence 51945, A
106	38	39.6	362	2	US-09-270-767-43340	Sequence 43340, A	179	37	38.5	102	2	US-09-270-767-34363	Sequence 34363, A
107	38	39.6	410	2	US-09-252-991A-26654	Sequence 26654, A	180	37	38.5	102	2	US-09-270-767-49580	Sequence 49580, A
108	38	39.6	427	2	US-09-902-540-14252	Sequence 14252, A	181	37	38.5	150	2	US-09-252-991A-27397	Sequence 27397, A
109	38	39.6	445	2	US-09-252-991A-23505	Sequence 23505, A	182	37	38.5	197	2	US-09-486-147-40	Sequence 40, Appl
110	38	39.6	448	2	US-09-252-991A-27497	Sequence 27497, A	183	37	38.5	226	2	US-09-134-000C-5787	Sequence 5787, Ap
111	38	39.6	478	2	US-09-252-991A-31455	Sequence 31455, A	184	37	38.5	237	2	US-09-902-540-13535	Sequence 13535, A
112	38	39.6	493	2	US-09-489-039A-8255	Sequence 8255, Ap	185	37	38.5	262	2	US-09-270-767-57404	Sequence 57404, A
113	38	39.6	529	2	US-09-252-991A-31840	Sequence 31840, A	186	37	38.5	269	2	US-09-252-991A-25667	Sequence 25667, A
114	38	39.6	531	2	US-09-328-352-6177	Sequence 6177, Ap	187	37	38.5	280	2	US-09-252-991A-20783	Sequence 20783, A
115	38	39.6	585	2	US-09-489-039A-13868	Sequence 13868, A	188	37	38.5	294	2	US-09-270-767-42136	Sequence 42136, A
116	38	39.6	597	2	US-08-704-966-2	Sequence 2, Appli	189	37	38.5	327	2	US-09-252-991A-30065	Sequence 30065, A
117	38	39.6	597	2	US-08-705-438-2	Sequence 4, Appli	190	37	38.5	332	2	US-09-134-000C-6066	Sequence 6066, Ap
118	38	39.6	610	2	US-08-704-966-4	Sequence 4, Appli	191	37	38.5	347	2	US-09-489-039A-8989	Sequence 8989, Ap
119	38	39.6	610	2	US-08-705-438-4	Sequence 4, Appli	192	37	38.5	363	2	US-08-096-181A-10	Sequence 10, Appl
120	38	39.6	613	2	US-09-252-991A-23154	Sequence 23154, A	193	37	38.5	363	4	PCT-US94-08326-10	Sequence 10, Appl
121	38	39.6	644	1	US-07-828-788A-4	Sequence 4, Appli	194	37	38.5	370	2	US-09-134-000C-5050	Sequence 5050, Ap
122	38	39.6	644	1	US-08-072-281-2	Sequence 2, Appli	195	37	38.5	373	2	US-09-602-777A-276	Sequence 276, App
123	38	39.6	644	1	US-08-295-060-2	Sequence 2, Appli	196	37	38.5	373	2	US-09-602-777A-386	Sequence 386, App
124	38	39.6	644	1	US-08-759-446-2	Sequence 2, Appli	197	37	38.5	390	2	US-09-943-313B-8	Sequence 8, Appli
125	38	39.6	644	1	US-09-027-998A-2	Sequence 2, Appli	198	37	38.5	453	2	US-09-489-039A-10870	Sequence 10870, A
126	38	39.6	644	2	US-09-696-348-2	Sequence 2, Appli	199	37	38.5	493	2	US-09-328-352-5214	Sequence 5214, Ap
127	38	39.6	644	2	US-10-222-441-2	Sequence 2, Appli	200	37	38.5	522	2	US-09-075-272-3	Sequence 3, Appli
128	38	39.6	644	2	PCT-US92-11337-4	Sequence 4, Appli	201	37	38.5	539	2	US-09-949-016-6363	Sequence 6363, Ap
129	38	39.6	645	1	US-08-295-060-4	Sequence 4, Appli	202	37	38.5	556	2	US-09-489-039A-9666	Sequence 9666, Ap
130	38	39.6	651	2	US-08-996-441B-52	Sequence 52, Appl	203	37	38.5	580	1	US-08-420-235B-15	Sequence 15, Appl
131	38	39.6	651	2	US-08-996-441B-56	Sequence 56, Appl	204	37	38.5	580	2	US-08-793-624-15	Sequence 15, Appl
132	38	39.6	651	2	US-08-996-441B-58	Sequence 58, Appl	205	37	38.5	580	2	US-09-252-991A-26637	Sequence 26637, A
133	38	39.6	651	2	US-08-993-722A-52	Sequence 52, Appl	206	37	38.5	580	4	PCT-US95-10194-15	Sequence 15, Appl
134	38	39.6	651	2	US-08-993-722A-56	Sequence 56, Appl	207	37	38.5	588	2	US-09-252-991A-31356	Sequence 31356, A
135	38	39.6	651	2	US-08-993-722A-58	Sequence 58, Appl	208	37	38.5	596	2	US-10-104-047-2541	Sequence 2541, Ap
136	38	39.6	651	2	US-08-993-722A-58	Sequence 58, Appl	209	37	38.5	622	2	US-09-075-272-2	Sequence 2, Appli
137	38	39.6	651	2	US-08-993-170A-52	Sequence 52, Appl	210	37	38.5	673	2	US-09-711-164-328	Sequence 328, App
138	38	39.6	651	2	US-08-993-170A-56	Sequence 56, Appl	211	37	38.5	686	2	US-09-949-016-11203	Sequence 11203, A
139	38	39.6	651	2	US-08-993-170A-58	Sequence 58, Appl	212	37	38.5	823	2	US-10-101-464A-923	Sequence 923, App
140	38	39.6	651	2	US-08-993-775B-52	Sequence 52, Appl	213	37	38.5	825	2	US-10-101-464A-922	Sequence 922, App
141	38	39.6	651	2	US-08-993-775B-56	Sequence 56, Appl	214	37	38.5	869	2	US-09-134-000C-5052	Sequence 5052, Ap
142	38	39.6	651	2	US-08-993-775B-58	Sequence 58, Appl	215	37	38.5	891	2	US-09-252-991A-25595	Sequence 25595, A
143	38	39.6	651	2	US-09-427-770-52	Sequence 52, Appl	216	37	38.5	920	2	US-10-104-047-2574	Sequence 2574, Ap
144	38	39.6	651	2	US-09-427-770-56	Sequence 56, Appl	217	37	38.5	933	1	US-08-370-193A-8	Sequence 8, Appli
145	38	39.6	651	2	US-09-427-770-58	Sequence 58, Appl	218	37	38.5	933	2	US-09-271-438A-9	Sequence 9, Appli
146	38	39.6	651	2	US-09-427-769-52	Sequence 52, Appl	219	37	38.5	933	2	US-10-078-107-5	Sequence 5, Appli
147	38	39.6	651	2	US-09-427-769-56	Sequence 56, Appl	220	37	38.5	935	2	US-09-271-438A-3	Sequence 3, Appli
148	38	39.6	651	2	US-08-996-441B-113	Sequence 113, App	221	37	38.5	935	2	US-09-271-438A-8	Sequence 8, Appli
149	38	39.6	652	2	US-08-996-441B-113	Sequence 113, App	222	37	38.5	935	2	US-10-078-107-1	Sequence 1, Appli
150	38	39.6	652	2	US-08-993-722A-113	Sequence 113, App	223	37	38.5	935	2	US-10-077-751-1	Sequence 1, Appli
151	38	39.6	652	2	US-08-993-170A-113	Sequence 113, App	224	37	38.5	938	2	US-09-489-039A-13504	Sequence 13504, A
152	38	39.6	652	2	US-08-993-775B-113	Sequence 113, App	225	37	38.5	1160	2	US-09-328-352-6826	Sequence 6826, Ap
153	38	39.6	652	2	US-09-427-770-113	Sequence 113, App	226	37	38.5	1189	2	US-09-489-039A-13776	Sequence 13776, A
154	38	39.6	652	2	US-09-427-769-113	Sequence 113, App	227	37	38.5	1205	2	US-09-330-330-1	Sequence 1, Appli
155	38	39.6	684	2	US-10-104-047-2418	Sequence 2418, Ap	228	37	38.5	1237	2	US-09-862-027-78	Sequence 78, Appl
156	38	39.6	687	2	US-09-489-039A-9868	Sequence 9868, Ap	229	37	38.5	42	2	US-09-270-767-40092	Sequence 40092, A
157	38	39.6	697	2	US-09-134-000C-5916	Sequence 5916, Ap	230	36.5	38.0	42	2	US-09-270-767-55308	Sequence 55308, A
158	38	39.6	742	2	US-09-489-039A-7998	Sequence 7998, Ap	231	36.5	38.0	42	2	US-09-605-703B-2204	Sequence 2204, Ap
159	38	39.6	749	2	US-09-562-737-97	Sequence 97, Appl	232	36.5	38.0	197	2	US-09-902-540-9704	Sequence 9704, Ap
160	38	39.6	853	2	US-09-354-129-10	Sequence 10, Appl	233	36.5	38.0	209	2	US-09-252-991A-24860	Sequence 24860, A
161	38	39.6	853	2	US-09-504-357-10	Sequence 10, Appl	234	36.5	38.0	412	2	US-09-647-540A-8	Sequence 8, Appli
162	38	39.6	937	2	US-09-543-681A-5513	Sequence 5513, Ap	235	36.5	38.0	432	2	US-09-647-540A-9	Sequence 9, Appli
163	38	39.6	1044	2	US-09-893-371A-1	Sequence 1, Appli	236	36.5	38.0	432	2	US-10-119-600-8	Sequence 8, Appli
164	38	39.6	1092	2	US-09-134-000C-6291	Sequence 6291, Ap	237	36.5	38.0	432	2	US-10-119-600-9	Sequence 9, Appli
165	38	39.6	1197	2	US-09-252-991A-30833	Sequence 30833, A	238	36.5	38.0	432	2	US-10-119-651-9	Sequence 9, Appli
166	38	39.6	1316	2	US-09-538-396-2	Sequence 2, Appli	239	36.5	38.0	432	2	US-09-252-991A-29979	Sequence 29979, A
167	38	39.6	1964	2	US-09-467-997-1	Sequence 1, Appli	240	36.5	38.0	580	2	US-09-418-710-13	Sequence 13, Appl
168	38	39.6	3571	2	US-09-911-842A-2	Sequence 2, Appli	241	36.5	38.0	1878	2	US-09-839-479-13	Sequence 13, Appl
169	38	39.6	4019	2	US-09-854-133-445	Sequence 425, App	242	36.5	38.0	1878	2	US-09-621-976-6389	Sequence 6389, Ap
170	37.5	39.1	476	2	US-10-104-047-3040	Sequence 3040, Ap	243	36.5	38.0	51	2	US-09-248-796A-26925	Sequence 26925, A
171	37.5	39.1	504	2	US-10-104-047-2889	Sequence 2889, Ap	244	36	37.5	68	2	US-09-489-039A-9120	Sequence 9120, Ap
172	37.5	39.1	770	2	US-09-784-316-5	Sequence 5, Appli	245	36	37.5	73	2		
173	37.5	39.1	770	2	US-10-229-124-5	Sequence 5, Appli	246	36	37.5				

247	36	37.5	82	2	US-09-763-509-16	Sequence 16, Appl	320	36	37.5	776	2	US-09-698-341-2	Sequence 2, Appl
248	36	37.5	85	2	US-09-252-991A-29339	Sequence 29339, A	321	36	37.5	778	1	US-08-906-925-4	Sequence 4, Appl
249	36	37.5	87	2	US-08-803-603-2	Sequence 2, Appl	322	36	37.5	843	2	US-09-328-352-5294	Sequence 5294, Ap
250	36	37.5	87	2	US-09-763-509-8	Sequence 8, Appl	323	36	37.5	847	2	US-09-252-991A-19346	Sequence 19346, A
251	36	37.5	87	2	US-09-763-509-15	Sequence 15, Appl	324	36	37.5	910	2	US-08-937-685A-2	Sequence 2, Appl
252	36	37.5	92	2	US-09-252-991A-27067	Sequence 27067, A	325	36	37.5	910	2	US-09-086-436-31	Sequence 31, Appl
253	36	37.5	95	2	US-09-252-991A-23142	Sequence 23142, A	326	36	37.5	937	2	US-09-252-991A-19446	Sequence 19446, A
254	36	37.5	96	2	US-09-763-509-6	Sequence 6, Appl	327	36	37.5	1007	1	US-08-551-459-4	Sequence 4, Appl
255	36	37.5	139	2	US-10-194-975-83	Sequence 83, Appl	328	36	37.5	1282	2	US-09-134-000C-5785	Sequence 5785, Ap
256	36	37.5	139	2	US-09-949-016-10706	Sequence 10706, A	329	36	37.5	1300	2	US-09-698-341-3	Sequence 3, Appl
257	36	37.5	148	2	US-09-949-016-10707	Sequence 10707, A	330	36	37.5	1785	2	US-09-341-587-3	Sequence 3, Appl
258	36	37.5	148	2	US-10-104-047-3547	Sequence 3547, Ap	331	36	37.5	2942	2	US-09-902-540-9733	Sequence 9733, Ap
259	36	37.5	149	2	US-09-902-540-13310	Sequence 13310, A	332	36	37.5	3782	2	US-09-105-537-4	Sequence 4, Appl
260	36	37.5	155	2	US-09-489-039A-9198	Sequence 9198, Ap	333	36	37.5	15281	1	US-08-471-119A-2	Sequence 2, Appl
261	36	37.5	162	2	US-10-104-047-3881	Sequence 3881, Ap	334	35.5	37.0	73	2	US-10-101-464A-482	Sequence 482, App
262	36	37.5	167	2	US-09-438-185A-132	Sequence 132, App	335	35.5	37.0	91	2	US-09-248-796A-14960	Sequence 14960, A
263	36	37.5	170	2	US-09-252-991A-25889	Sequence 25889, A	336	35.5	37.0	99	2	US-09-489-039A-7903	Sequence 7903, Ap
264	36	37.5	202	2	US-09-252-991A-26102	Sequence 26102, A	337	35.5	37.0	125	2	US-09-107-532A-6496	Sequence 6496, Ap
265	36	37.5	208	2	US-09-489-039A-12178	Sequence 12178, A	338	35.5	37.0	212	2	US-09-902-540-14657	Sequence 14657, A
266	36	37.5	211	2	US-09-252-991A-27027	Sequence 27027, A	339	35.5	37.0	540	2	US-09-902-540-9824	Sequence 9824, Ap
267	36	37.5	240	2	US-08-824-692-23	Sequence 23, Appl	340	35.5	37.0	703	2	US-09-252-991A-31954	Sequence 31954, A
268	36	37.5	253	2	US-09-188-930-128	Sequence 128, App	341	35.5	37.0	749	2	US-09-562-737-96	Sequence 96, Appl
269	36	37.5	253	2	US-09-312-283C-128	Sequence 128, App	342	35.5	37.0	788	2	US-09-248-796A-20511	Sequence 20511, A
270	36	37.5	255	2	US-09-540-236-2831	Sequence 2831, App	343	35.5	37.0	861	2	US-09-784-316-2	Sequence 2, Appl
271	36	37.5	281	2	US-09-252-991A-25471	Sequence 25471, A	344	35.5	37.0	861	2	US-10-229-124-2	Sequence 2, Appl
272	36	37.5	302	2	US-09-198-452A-18	Sequence 18, Appl	345	35.5	37.0	951	2	US-09-328-352-4167	Sequence 4167, Ap
273	36	37.5	302	2	US-09-438-185A-9	Sequence 9, Appl	346	35.5	37.0	1336	1	US-08-231-193A-58	Sequence 58, Appl
274	36	37.5	326	2	US-09-787-192-11	Sequence 11, Appl	347	35.5	37.0	1336	1	US-08-486-273A-58	Sequence 58, Appl
275	36	37.5	327	2	US-09-902-540-15162	Sequence 15162, A	348	35.5	37.0	1336	2	US-08-940-086A-58	Sequence 58, Appl
276	36	37.5	328	2	US-09-787-192-2	Sequence 2, Appl	349	35.5	37.0	1336	2	US-08-940-035A-58	Sequence 58, Appl
277	36	37.5	365	2	US-10-104-047-3375	Sequence 3375, Ap	350	35.5	37.0	1336	2	US-08-935-105A-58	Sequence 58, Appl
278	36	37.5	373	2	US-09-149-476-374	Sequence 374, App	351	35.5	37.0	1336	2	US-09-648-797-58	Sequence 58, Appl
279	36	37.5	380	2	US-09-252-991A-24774	Sequence 24774, A	352	35.5	37.0	1336	2	US-09-386-123-58	Sequence 58, Appl
280	36	37.5	385	2	US-09-328-352-6804	Sequence 6804, Ap	353	35.5	37.0	1336	2	US-10-038-937-58	Sequence 58, Appl
281	36	37.5	396	2	US-09-198-452A-147	Sequence 147, App	354	35.5	37.0	1336	2	US-10-007-747-58	Sequence 58, Appl
282	36	37.5	405	2	US-09-252-991A-26539	Sequence 26539, A	355	35.5	37.0	1336	2	US-09-945-901-58	Sequence 58, Appl
283	36	37.5	415	2	US-09-252-991A-23751	Sequence 23751, A	356	35.5	37.0	1345	2	US-09-949-016-11209	Sequence 11209, A
284	36	37.5	434	2	US-09-902-540-12368	Sequence 12368, A	357	35	36.5	20	2	US-09-962-756-1068	Sequence 1068, Ap
285	36	37.5	437	2	US-09-039-064-3	Sequence 3, Appl	358	35	36.5	56	1	US-08-162-102C-39	Sequence 39, Appl
286	36	37.5	448	2	US-09-902-540-10764	Sequence 10764, A	359	35	36.5	58	2	US-09-621-976-7056	Sequence 7056, Ap
287	36	37.5	454	2	US-09-450-209-16	Sequence 16, Appl	360	35	36.5	66	2	US-09-605-703B-1036	Sequence 1036, Ap
288	36	37.5	485	2	US-09-320-878-10	Sequence 10, Appl	361	35	36.5	73	2	US-09-605-703B-1038	Sequence 1038, Ap
289	36	37.5	485	2	US-09-105-537-10	Sequence 10, Appl	362	35	36.5	77	2	US-09-248-796A-27926	Sequence 27926, A
290	36	37.5	485	2	US-09-141-908-23	Sequence 23, Appl	363	35	36.5	81	2	US-09-311-021-12	Sequence 12, Appl
291	36	37.5	485	2	US-09-657-440-10	Sequence 10, Appl	364	35	36.5	96	2	US-09-252-991A-20300	Sequence 20300, A
292	36	37.5	485	2	US-09-793-708-10	Sequence 10, Appl	365	35	36.5	100	2	US-09-248-796A-21368	Sequence 21368, A
293	36	37.5	493	2	US-09-543-681A-7006	Sequence 7006, Ap	366	35	36.5	104	2	US-08-767-128-12	Sequence 12, Appl
294	36	37.5	510	2	US-09-489-039A-10171	Sequence 10171, A	367	35	36.5	104	2	US-08-767-128-14	Sequence 14, Appl
295	36	37.5	516	2	US-09-902-540-12648	Sequence 12648, A	368	35	36.5	104	2	US-08-767-128-16	Sequence 16, Appl
296	36	37.5	520	2	US-09-252-991A-17942	Sequence 17942, A	369	35	36.5	106	1	US-08-561-521-5	Sequence 5, Appl
297	36	37.5	550	2	US-09-252-991A-19086	Sequence 19086, A	370	35	36.5	106	4	PCT-US95-01219-5	Sequence 5, Appl
298	36	37.5	553	1	US-08-565-386-6	Sequence 6, Appl	371	35	36.5	107	2	US-08-767-128-30	Sequence 30, Appl
299	36	37.5	561	2	US-09-949-016-11704	Sequence 11704, A	372	35	36.5	107	2	US-08-767-128-34	Sequence 34, Appl
300	36	37.5	568	2	US-09-160-065-2	Sequence 2, Appl	373	35	36.5	109	1	US-08-162-102C-24	Sequence 24, Appl
301	36	37.5	589	2	US-09-252-991A-18210	Sequence 18210, A	374	35	36.5	109	4	US-08-920-100B-24	Sequence 24, Appl
302	36	37.5	617	2	US-09-252-991A-21113	Sequence 21113, A	375	35	36.5	109	4	PCT-US93-08786-24	Sequence 24, Appl
303	36	37.5	632	2	US-09-853-533A-8	Sequence 8, Appl	376	35	36.5	111	2	US-09-489-039A-12797	Sequence 12797, A
304	36	37.5	642	2	US-09-543-681A-7930	Sequence 7930, Ap	377	35	36.5	113	2	US-08-983-607-25	Sequence 25, Appl
305	36	37.5	649	2	US-09-936-441B-109	Sequence 109, App	378	35	36.5	126	1	US-08-561-521-2	Sequence 2, Appl
306	36	37.5	649	2	US-08-993-722A-109	Sequence 109, App	379	35	36.5	126	1	US-08-561-521-15	Sequence 15, Appl
307	36	37.5	649	2	US-08-993-170A-109	Sequence 109, App	380	35	36.5	126	2	US-08-483-533-31	Sequence 31, Appl
308	36	37.5	649	2	US-08-993-775B-109	Sequence 109, App	381	35	36.5	126	2	US-09-283-471A-31	Sequence 31, Appl
309	36	37.5	649	2	US-09-295-186-18	Sequence 18, Appl	382	35	36.5	126	4	PCT-US95-01219-2	Sequence 2, Appl
310	36	37.5	649	2	US-09-427-770-109	Sequence 109, App	383	35	36.5	126	4	PCT-US95-01219-15	Sequence 15, Appl
311	36	37.5	649	2	US-09-427-769-109	Sequence 109, App	384	35	36.5	131	2	US-09-328-352-4847	Sequence 4847, Ap
312	36	37.5	688	2	US-09-248-796A-17856	Sequence 17856, A	385	35	36.5	156	2	US-09-252-991A-31849	Sequence 31849, A
313	36	37.5	692	2	US-09-758-759-79	Sequence 79, Appl	386	35	36.5	163	2	US-08-858-207A-451	Sequence 451, App
314	36	37.5	693	2	US-09-252-991A-20348	Sequence 20348, A	387	35	36.5	172	2	US-09-252-991A-25305	Sequence 25305, A
315	36	37.5	731	2	US-09-270-767-42057	Sequence 42057, A	388	35	36.5	175	2	US-09-605-703B-2624	Sequence 2624, Ap
316	36	37.5	738	2	US-08-989-385-1	Sequence 1, Appl	389	35	36.5	178	2	US-08-705-875A-8	Sequence 8, Appl
317	36	37.5	738	2	US-09-593-826-1	Sequence 1, Appl	390	35	36.5	178	2	US-09-220-731-23	Sequence 23, Appl
318	36	37.5	738	2	US-09-949-016-6020	Sequence 6020, Ap	391	35	36.5	178	2	US-09-220-731-24	Sequence 24, Appl
319	36	37.5	753	2	US-09-949-016-11542	Sequence 11542, A	392	35	36.5	178	2	US-09-242-999-8	Sequence 8, Appl

393	35	36.5	178	2	US-09-242-999-24	Sequence 24, Appl	466	35	36.5	415	2	US-09-252-991A-31846	Sequence 31846, A
394	35	36.5	179	1	US-08-752-132-2	Sequence 2, Appl	467	35	36.5	425	2	US-09-489-039A-8386	Sequence 8386, Ap
395	35	36.5	179	2	US-08-483-533-33	Sequence 33, Appl	468	35	36.5	426	2	US-09-538-032-568	Sequence 568, App
396	35	36.5	179	2	US-09-283-471A-33	Sequence 33, Appl	469	35	36.5	450	2	US-09-252-991A-16659	Sequence 16659, A
397	35	36.5	180	2	US-10-104-047-3533	Sequence 33, Appl	470	35	36.5	481	2	US-09-949-016-9549	Sequence 9549, Ap
398	35	36.5	183	2	US-09-902-540-14351	Sequence 14351, A	471	35	36.5	482	2	US-09-248-796A-17800	Sequence 17800, A
399	35	36.5	185	2	US-09-252-991A-20667	Sequence 20667, A	472	35	36.5	489	1	US-08-095-726-8	Sequence 8, Appl
400	35	36.5	197	2	US-09-252-991A-30009	Sequence 30009, A	473	35	36.5	489	1	US-08-095-726-8	Sequence 8, Appl
401	35	36.5	197	2	US-09-252-991A-22880	Sequence 22880, A	474	35	36.5	489	1	US-08-095-726-8	Sequence 8, Appl
402	35	36.5	203	2	US-09-583-110-4477	Sequence 4477, Ap	475	35	36.5	489	1	US-08-096-043-10	Sequence 10, Appl
403	35	36.5	212	2	US-09-302-769-12	Sequence 12, Appl	476	35	36.5	489	1	US-08-096-043-10	Sequence 10, Appl
404	35	36.5	212	2	US-09-399-913-63	Sequence 63, Appl	477	35	36.5	489	1	US-08-096-043-10	Sequence 10, Appl
405	35	36.5	212	2	US-09-350-614-63	Sequence 63, Appl	478	35	36.5	493	2	US-09-252-991A-31287	Sequence 31287, A
406	35	36.5	212	2	US-08-962-560C-12	Sequence 12, Appl	479	35	36.5	496	2	US-09-252-991A-21949	Sequence 21949, A
407	35	36.5	216	2	US-08-824-692-24	Sequence 24, Appl	480	35	36.5	500	1	US-08-307-499-20	Sequence 20, Appl
408	35	36.5	221	2	US-09-252-991A-18874	Sequence 18874, A	481	35	36.5	500	2	US-09-299-268-20	Sequence 20, Appl
409	35	36.5	228	2	US-10-104-047-3125	Sequence 3125, Ap	482	35	36.5	506	2	US-09-134-001C-4049	Sequence 4049, Ap
410	35	36.5	245	2	US-08-483-533-42	Sequence 42, Appl	483	35	36.5	511	2	US-09-902-540-15937	Sequence 15937, A
411	35	36.5	245	2	US-09-283-471A-42	Sequence 42, Appl	484	35	36.5	513	2	US-09-252-991A-27380	Sequence 27380, A
412	35	36.5	249	2	US-09-649-747A-14	Sequence 14, Appl	485	35	36.5	516	2	US-09-291-170A-1	Sequence 1, Appl
413	35	36.5	249	2	US-09-649-747A-15	Sequence 15, Appl	486	35	36.5	516	2	US-09-724-884-1	Sequence 1, Appl
414	35	36.5	252	2	US-08-483-533-43	Sequence 43, Appl	487	35	36.5	517	2	US-09-724-884-1	Sequence 1, Appl
415	35	36.5	252	2	US-09-283-471A-43	Sequence 43, Appl	488	35	36.5	517	2	US-09-673-522-1	Sequence 1, Appl
416	35	36.5	254	2	US-09-328-352-7193	Sequence 7193, Ap	489	35	36.5	541	2	US-09-252-991A-19837	Sequence 19837, A
417	35	36.5	254	2	US-09-808-387-32	Sequence 32, Appl	490	35	36.5	545	2	US-10-237-551-121	Sequence 121, App
418	35	36.5	258	2	US-08-483-533-26	Sequence 26, Appl	491	35	36.5	545	2	US-10-237-551-157	Sequence 157, App
419	35	36.5	258	2	US-09-283-471A-26	Sequence 26, Appl	492	35	36.5	547	2	US-10-237-551-215	Sequence 215, App
420	35	36.5	258	2	US-09-283-471A-26	Sequence 26, Appl	493	35	36.5	547	2	US-10-237-551-216	Sequence 216, App
421	35	36.5	263	4	PCT-US91-06532-2	Sequence 2, Appl	494	35	36.5	566	2	US-09-902-540-10009	Sequence 10009, A
422	35	36.5	264	2	US-08-483-533-40	Sequence 40, Appl	495	35	36.5	568	2	US-09-949-016-7970	Sequence 7970, Ap
423	35	36.5	264	2	US-09-283-471A-40	Sequence 40, Appl	496	35	36.5	572	2	US-09-487-558B-246	Sequence 246, App
424	35	36.5	279	2	US-09-902-540-12548	Sequence 12548, A	497	35	36.5	581	2	US-09-649-747A-13	Sequence 13, Appl
425	35	36.5	282	2	US-09-602-787A-388	Sequence 388, App	498	35	36.5	581	2	US-09-649-747A-21	Sequence 21, Appl
426	35	36.5	294	2	US-09-252-991A-26953	Sequence 26953, A	499	35	36.5	584	2	US-09-842-307-2	Sequence 2, Appl
427	35	36.5	295	2	US-09-252-991A-19985	Sequence 19985, A	500	35	36.5	610	2	US-09-709-457-2	Sequence 2, Appl
428	35	36.5	301	2	US-09-248-796A-14373	Sequence 14373, A	501	35	36.5	630	2	US-09-252-991A-19702	Sequence 19702, A
429	35	36.5	302	2	US-09-220-731-26	Sequence 26, Appl	502	35	36.5	642	2	US-09-902-540-16286	Sequence 16286, A
430	35	36.5	302	2	US-09-242-999-22	Sequence 22, Appl	503	35	36.5	710	2	US-09-328-352-5900	Sequence 5900, Ap
431	35	36.5	302	2	US-09-252-991A-16900	Sequence 16900, A	504	35	36.5	714	2	US-09-252-991A-31209	Sequence 31209, A
432	35	36.5	303	2	US-09-248-796A-14375	Sequence 14375, A	505	35	36.5	726	2	US-09-252-991A-20675	Sequence 20675, A
433	35	36.5	305	2	US-09-270-767-46577	Sequence 46577, A	506	35	36.5	747	2	US-09-134-000C-6641	Sequence 6641, Ap
434	35	36.5	308	2	US-08-705-875A-10	Sequence 10, Appl	507	35	36.5	771	2	US-09-214-916A-2	Sequence 2, Appl
435	35	36.5	308	2	US-09-242-999-10	Sequence 10, Appl	508	35	36.5	771	2	US-09-214-916A-4	Sequence 4, Appl
436	35	36.5	313	2	US-09-438-185A-1056	Sequence 1056, Ap	509	35	36.5	771	2	US-09-248-796A-20725	Sequence 20725, A
437	35	36.5	315	2	US-09-187-906-19	Sequence 19, Appl	510	35	36.5	773	2	US-09-585-858-35	Sequence 35, Appl
438	35	36.5	315	2	US-09-489-407-19	Sequence 19, Appl	511	35	36.5	773	2	US-10-270-878-35	Sequence 35, Appl
439	35	36.5	320	2	US-09-198-452A-1130	Sequence 1130, Ap	512	35	36.5	774	2	US-08-902-632-2	Sequence 2, Appl
440	35	36.5	323	2	US-09-489-039A-13655	Sequence 13655, A	513	35	36.5	774	2	US-09-073-354-1	Sequence 1, Appl
441	35	36.5	330	2	US-09-583-110-3623	Sequence 3623, Ap	514	35	36.5	774	2	US-08-656-005A-1	Sequence 1, Appl
442	35	36.5	333	2	US-09-252-991A-17113	Sequence 17113, A	515	35	36.5	774	2	US-09-073-259-1	Sequence 1, Appl
443	35	36.5	351	2	US-09-399-913-61	Sequence 61, Appl	516	35	36.5	774	2	US-09-363-095-1	Sequence 1, Appl
444	35	36.5	351	2	US-09-489-039A-9415	Sequence 9415, Ap	517	35	36.5	774	2	US-09-418-027-1	Sequence 1, Appl
445	35	36.5	351	2	US-09-350-614-61	Sequence 61, Appl	518	35	36.5	776	2	US-09-165-396-3	Sequence 3, Appl
446	35	36.5	358	2	US-10-104-047-2376	Sequence 2376, Ap	519	35	36.5	812	2	US-09-328-352-5434	Sequence 5434, Ap
447	35	36.5	361	2	US-09-489-039A-12917	Sequence 12917, A	520	35	36.5	862	2	US-09-252-991A-19318	Sequence 19318, A
448	35	36.5	368	2	US-09-902-540-15739	Sequence 15739, A	521	35	36.5	864	1	US-08-620-694A-2	Sequence 2, Appl
449	35	36.5	369	2	US-09-328-352-5957	Sequence 5957, Ap	522	35	36.5	864	2	US-09-022-255-2	Sequence 2, Appl
450	35	36.5	370	2	US-09-134-001C-3403	Sequence 3403, Ap	523	35	36.5	864	2	US-09-022-696-2	Sequence 2, Appl
451	35	36.5	370	2	US-09-710-279-696	Sequence 696, App	524	35	36.5	864	2	US-08-978-773-2	Sequence 2, Appl
452	35	36.5	372	2	US-09-710-279-1328	Sequence 1328, Ap	525	35	36.5	864	2	US-09-022-253-2	Sequence 2, Appl
453	35	36.5	372	2	US-09-406-045-1	Sequence 1, Appl	526	35	36.5	864	2	US-09-022-260-2	Sequence 2, Appl
454	35	36.5	380	2	US-09-248-796A-16419	Sequence 16419, A	527	35	36.5	864	2	US-09-022-257-2	Sequence 2, Appl
455	35	36.5	384	2	US-09-252-991A-18786	Sequence 18786, A	528	35	36.5	864	2	US-09-022-257-2	Sequence 2, Appl
456	35	36.5	390	2	US-09-252-991A-18471	Sequence 18471, A	529	35	36.5	864	2	US-09-549-679-2	Sequence 2, Appl
457	35	36.5	395	2	US-09-673-395A-612	Sequence 612, App	530	35	36.5	895	2	US-09-614-912-194	Sequence 194, App
458	35	36.5	396	3	US-08-229-284A-2	Sequence 2, Appl	531	35	36.5	912	2	US-09-540-236-1960	Sequence 1960, Ap
459	35	36.5	399	2	US-09-270-767-45921	Sequence 45921, A	532	35	36.5	967	2	US-09-540-236-2449	Sequence 2449, Ap
460	35	36.5	400	2	US-09-220-528-63	Sequence 63, Appl	533	35	36.5	1019	1	US-08-271-364A-7	Sequence 7, Appl
461	35	36.5	400	2	US-09-187-906-21	Sequence 21, Appl	534	35	36.5	1019	1	US-08-222-715B-26	Sequence 26, Appl
462	35	36.5	400	2	US-09-949-016-9079	Sequence 9079, Ap	535	35	36.5	1109	2	US-09-949-016-6671	Sequence 10771, A
463	35	36.5	400	2	US-09-489-407-21	Sequence 21, Appl	536	35	36.5	1203	2	US-09-949-016-6671	Sequence 6615, Ap
464	35	36.5	401	2	US-09-949-016-6223	Sequence 8223, Ap	537	35	36.5	1290	1	US-08-138-641-2	Sequence 2, Appl
465	35	36.5	412	2	US-09-252-991A-29286	Sequence 29286, A	538	35	36.5	1290	1	US-08-138-133-2	Sequence 2, Appl

539	35	36.5	1290	2	US-09-538-092-956	Sequence 956, App	612	34	35.4	189	2	US-09-513-999C-7827	Sequence 7827, Ap
540	35	36.5	1312	2	US-09-949-016-10141	Sequence 10141, A	613	34	35.4	189	2	US-09-513-999C-7828	Sequence 7828, Ap
541	35	36.5	2190	2	US-09-252-991A-25754	Sequence 25754, A	614	34	35.4	196	2	US-09-252-991A-22525	Sequence 22525, A
542	35	36.5	2636	2	US-09-252-991A-25753	Sequence 25753, A	615	34	35.4	197	2	US-09-270-767-33420	Sequence 33420, A
543	34.5	35.9	20	2	US-09-690-454-100	Sequence 100, App	616	34	35.4	197	2	US-09-270-767-48637	Sequence 48637, A
544	34.5	35.9	38	2	US-09-270-767-35903	Sequence 35903, App	617	34	35.4	204	2	US-08-965-056-105	Sequence 105, App
545	34.5	35.9	38	2	US-09-270-767-31120	Sequence 31120, A	618	34	35.4	208	2	US-09-489-039A-11072	Sequence 11072, A
546	34.5	35.9	66	2	US-09-227-357-305	Sequence 205, App	619	34	35.4	211	2	US-09-252-991A-17038	Sequence 17038, A
547	34.5	35.9	66	2	US-09-973-278-246	Sequence 246, App	620	34	35.4	211	2	US-09-489-039A-7444	Sequence 7444, Ap
548	34.5	35.9	93	2	US-10-101-464A-779	Sequence 779, App	621	34	35.4	216	2	US-09-543-681A-5729	Sequence 5729, Ap
549	34.5	35.9	121	2	US-09-489-039A-12666	Sequence 12666, A	622	34	35.4	218	2	US-09-893-737-28	Sequence 28, Appl
550	34.5	35.9	162	2	US-10-104-047-3256	Sequence 3256, App	623	34	35.4	225	2	US-09-252-991A-32897	Sequence 32897, A
551	34.5	35.9	180	2	US-09-328-352-3483	Sequence 3483, Ap	624	34	35.4	229	2	US-09-252-991A-23807	Sequence 23807, A
552	34.5	35.9	196	2	US-09-690-454-94	Sequence 94, Appl	625	34	35.4	233	2	US-10-305-770B-2	Sequence 2, Appl
553	34.5	35.9	257	2	US-09-902-540-12292	Sequence 12292, A	626	34	35.4	237	2	US-09-270-767-39303	Sequence 39303, A
554	34.5	35.9	330	2	US-09-949-016-9848	Sequence 9848, Ap	627	34	35.4	237	2	US-09-270-767-54520	Sequence 54520, A
555	34.5	35.9	425	2	US-09-252-991A-17186	Sequence 17186, A	628	34	35.4	246	2	US-09-902-540-13603	Sequence 13603, A
556	34.5	35.9	434	1	US-08-484-575A-20	Sequence 20, Appl	629	34	35.4	251	2	US-09-252-991A-24677	Sequence 24677, A
557	34.5	35.9	434	2	US-08-477-459-20	Sequence 20, Appl	630	34	35.4	251	2	US-09-252-991A-28124	Sequence 28124, A
558	34.5	35.9	434	2	US-08-479-869-20	Sequence 20, Appl	631	34	35.4	251	2	US-09-107-532A-5498	Sequence 5498, Ap
559	34.5	35.9	434	2	US-08-486-414-46	Sequence 46, Appl	632	34	35.4	255	2	US-09-107-532A-5223	Sequence 5223, Ap
560	34.5	35.9	434	2	US-09-993-777-11	Sequence 11, Appl	633	34	35.4	257	1	US-08-467-265-16	Sequence 16, Appl
561	34.5	35.9	434	2	US-09-993-777-67	Sequence 67, Appl	634	34	35.4	257	2	US-08-467-265-16	Sequence 16, Appl
562	34.5	35.9	434	2	US-09-881-457A-4	Sequence 4, Appl	635	34	35.4	257	2	US-09-407-891-16	Sequence 16, Appl
563	34.5	35.9	434	4	PCT-US94-01826A-20	Sequence 20, Appl	636	34	35.4	257	2	US-09-375-907-5	Sequence 5, Appl
564	34.5	35.9	434	4	PCT-US94-02252A-20	Sequence 20, Appl	637	34	35.4	261	2	US-09-602-565-34	Sequence 34, Appl
565	34.5	35.9	434	4	PCT-US96-03916-11	Sequence 11, Appl	638	34	35.4	270	2	US-09-489-039A-11244	Sequence 11244, A
566	34.5	35.9	434	4	PCT-US96-03916-67	Sequence 67, Appl	639	34	35.4	276	2	US-09-252-991A-25159	Sequence 25159, A
567	34.5	35.9	478	1	US-08-951-148-7	Sequence 7, Appl	640	34	35.4	278	2	US-09-252-991A-22273	Sequence 22273, A
568	34.5	35.9	478	1	US-08-165-234-7	Sequence 7, Appl	641	34	35.4	285	2	US-09-369-48A-2	Sequence 2, Appl
569	34.5	35.9	478	2	US-09-274-570-7	Sequence 7, Appl	642	34	35.4	286	2	US-09-252-991A-23854	Sequence 23854, A
570	34.5	35.9	490	1	US-08-687-916-34	Sequence 24, Appl	643	34	35.4	286	2	US-09-543-681A-4840	Sequence 4840, Ap
571	34.5	35.9	490	2	US-09-138-614-24	Sequence 24, Appl	644	34	35.4	291	2	US-09-252-991A-19638	Sequence 19638, A
572	34.5	35.9	608	2	US-09-413-814-92	Sequence 92, Appl	645	34	35.4	302	2	US-09-489-039A-10944	Sequence 10944, A
573	34.5	35.9	624	2	US-09-489-039A-8321	Sequence 8321, Ap	646	34	35.4	303	2	US-09-107-532A-4923	Sequence 4923, Ap
574	34.5	35.9	1213	2	US-09-413-814-79	Sequence 79, Appl	647	34	35.4	305	2	US-09-674-529B-14	Sequence 14, Appl
575	34.5	35.9	1876	2	US-09-418-710-71	Sequence 71, Appl	648	34	35.4	308	2	US-09-347-803-12	Sequence 12, Appl
576	34.5	35.9	2183	2	US-09-839-479-70	Sequence 70, Appl	649	34	35.4	317	2	US-09-134-000C-5654	Sequence 5654, Ap
577	34.5	35.9	2183	2	US-08-746-111-5	Sequence 5, Appl	650	34	35.4	317	2	US-09-489-039A-9768	Sequence 9768, Ap
578	34	35.4	55	2	US-08-955-937A-9	Sequence 9, Appl	651	34	35.4	324	2	US-09-489-039A-10801	Sequence 10801, A
579	34	35.4	55	2	US-08-300-985A-9	Sequence 9, Appl	652	34	35.4	330	2	US-09-538-092-180	Sequence 180, App
580	34	35.4	56	2	US-08-936-165A-297	Sequence 297, App	653	34	35.4	330	2	US-09-902-540-15620	Sequence 15620, A
581	34	35.4	61	2	US-09-216-973B-122	Sequence 122, App	654	34	35.4	332	2	US-09-270-767-42978	Sequence 42978, A
582	34	35.4	69	2	US-09-621-976-4462	Sequence 4462, App	655	34	35.4	333	2	US-09-664-958-1	Sequence 1, Appl
583	34	35.4	73	2	US-09-149-076-365	Sequence 365, App	656	34	35.4	333	2	US-09-664-958-14	Sequence 14, Appl
584	34	35.4	78	2	US-09-489-039A-11104	Sequence 11104, A	657	34	35.4	334	2	US-09-438-185A-11	Sequence 11, Appl
585	34	35.4	79	2	US-09-513-999C-6001	Sequence 6001, Ap	658	34	35.4	340	2	US-09-142-593-1	Sequence 1, Appl
586	34	35.4	86	2	US-09-489-039A-12953	Sequence 12953, A	659	34	35.4	340	2	US-09-927-886-1	Sequence 1, Appl
587	34	35.4	95	2	US-10-194-975-59	Sequence 59, Appl	660	34	35.4	346	2	US-09-198-452A-60	Sequence 60, Appl
588	34	35.4	95	2	US-10-194-975-71	Sequence 71, Appl	661	34	35.4	351	1	US-08-324-483-2	Sequence 2, Appl
589	34	35.4	95	2	US-10-330-613A-47	Sequence 47, Appl	662	34	35.4	351	1	US-08-470-202-46	Sequence 46, Appl
590	34	35.4	95	2	US-10-330-613A-48	Sequence 48, Appl	663	34	35.4	351	1	US-08-471-770-46	Sequence 46, Appl
591	34	35.4	95	2	US-10-330-613A-51	Sequence 51, Appl	664	34	35.4	351	1	US-08-468-059-46	Sequence 46, Appl
592	34	35.4	95	2	US-10-330-613A-52	Sequence 52, Appl	665	34	35.4	351	2	US-09-109-916-46	Sequence 46, Appl
593	34	35.4	100	2	US-08-654-482-12	Sequence 12, Appl	666	34	35.4	351	2	US-09-886-156-46	Sequence 46, Appl
594	34	35.4	100	2	US-08-899-886-12	Sequence 12, Appl	667	34	35.4	351	2	US-09-886-149-46	Sequence 46, Appl
595	34	35.4	100	2	US-09-585-023-12	Sequence 12, Appl	668	34	35.4	351	2	US-09-886-150-46	Sequence 46, Appl
596	34	35.4	101	2	US-09-489-039A-8406	Sequence 8406, Ap	669	34	35.4	351	2	US-09-886-159-46	Sequence 46, Appl
597	34	35.4	107	2	US-10-330-613A-14	Sequence 14, Appl	670	34	35.4	351	2	US-10-326-090-46	Sequence 46, Appl
598	34	35.4	107	2	US-10-330-613A-18	Sequence 18, Appl	671	34	35.4	357	2	US-09-248-796A-20175	Sequence 20175, A
599	34	35.4	108	2	US-09-252-991A-28865	Sequence 28865, A	672	34	35.4	358	2	US-08-858-207A-398	Sequence 398, App
600	34	35.4	108	2	US-09-902-540-11349	Sequence 11349, A	673	34	35.4	375	2	US-09-252-991A-31128	Sequence 31128, A
601	34	35.4	135	2	US-09-252-991A-17289	Sequence 17289, A	674	34	35.4	376	2	US-09-252-991A-25057	Sequence 25057, A
602	34	35.4	139	2	US-09-621-976-6561	Sequence 6561, Ap	675	34	35.4	377	2	US-09-902-540-13007	Sequence 13007, A
603	34	35.4	146	1	US-08-394-021-10	Sequence 10, Appl	676	34	35.4	377	2	US-09-252-991A-29332	Sequence 29332, A
604	34	35.4	146	2	US-09-131-551-10	Sequence 10, Appl	677	34	35.4	383	2	US-09-603-208A-62	Sequence 62, Appl
605	34	35.4	146	2	US-10-000-321-10	Sequence 10, Appl	678	34	35.4	384	2	US-09-107-532A-7144	Sequence 7144, Ap
606	34	35.4	148	2	US-09-252-991A-32368	Sequence 32368, A	679	34	35.4	393	2	US-09-489-039A-7305	Sequence 7305, Ap
607	34	35.4	153	2	US-09-902-540-14298	Sequence 14298, A	680	34	35.4	398	2	US-09-543-681A-8342	Sequence 8342, Ap
608	34	35.4	165	2	US-09-252-991A-22416	Sequence 22416, A	681	34	35.4	407	2	US-09-252-991A-20436	Sequence 20436, A
609	34	35.4	170	2	US-09-252-991A-17086	Sequence 17086, A	682	34	35.4	412	2	US-09-252-991A-18513	Sequence 18513, A
610	34	35.4	170	2	US-09-902-540-11742	Sequence 11742, A	683	34	35.4	413	2	US-09-252-991A-24646	Sequence 24646, A
611	34	35.4	171	2	US-09-252-991A-30524	Sequence 30524, A	684	34	35.4	420	1	US-08-864-804-1	Sequence 1, Appl



685	34	35.4	422	2	US-09-252-991A-21608	Sequence 21608, A	758	34	35.4	898	1	US-08-966-145-2	Sequence 2, Appli
686	34	35.4	422	2	US-09-248-796A-14141	Sequence 14141, A	759	34	35.4	898	1	US-08-966-145-4	Sequence 4, Appli
687	34	35.4	424	2	US-09-252-991A-22384	Sequence 22384, A	760	34	35.4	920	1	US-08-101-593-2	Sequence 2, Appli
688	34	35.4	427	2	US-08-705-771-16	Sequence 16, Appl	761	34	35.4	920	1	US-08-101-593-4	Sequence 4, Appli
689	34	35.4	427	2	US-09-417-540-16	Sequence 16, Appl	762	34	35.4	935	2	US-09-477-962-107	Sequence 107, App
690	34	35.4	427	2	US-09-949-016-6096	Sequence 6096, Ap	763	34	35.4	964	2	US-09-543-681A-5035	Sequence 5035, Ap
691	34	35.4	431	2	US-09-252-991A-26751	Sequence 26751, A	764	34	35.4	1008	2	US-09-252-991A-24402	Sequence 24402, A
692	34	35.4	432	2	PCT-US95-04910-13	Sequence 13, Appl	765	34	35.4	998	2	US-09-252-991A-27063	Sequence 27063, A
693	34	35.4	442	2	US-09-949-016-7374	Sequence 7374, Ap	766	34	35.4	1022	1	US-08-271-364A-8	Sequence 8, Appli
694	34	35.4	458	2	US-09-252-991A-32991	Sequence 32991, A	767	34	35.4	1022	1	US-08-222-715B-27	Sequence 27, Appl
695	34	35.4	463	2	US-09-489-039A-8259	Sequence 8259, Ap	768	34	35.4	1049	2	US-10-018-730A-4	Sequence 4, Appli
696	34	35.4	465	2	US-09-489-039A-8545	Sequence 8545, Ap	769	34	35.4	1050	2	US-09-538-092-131	Sequence 131, App
697	34	35.4	472	2	US-09-252-991A-16723	Sequence 16723, A	770	34	35.4	1076	2	US-09-949-016-6610	Sequence 6610, Ap
698	34	35.4	475	2	US-09-949-016-11570	Sequence 11570, A	771	34	35.4	1090	2	US-09-866-510-14	Sequence 14, Appl
699	34	35.4	477	2	US-09-252-991A-29156	Sequence 29156, A	772	34	35.4	1100	2	US-09-949-016-7524	Sequence 7524, Ap
700	34	35.4	484	2	US-10-104-047-2240	Sequence 2240, Ap	773	34	35.4	1106	1	US-08-180-195-2	Sequence 2, Appli
701	34	35.4	485	2	US-09-438-185A-43	Sequence 43, Appl	774	34	35.4	1106	1	US-08-188-917-2	Sequence 2, Appli
702	34	35.4	491	2	US-09-328-352-5438	Sequence 5438, Ap	775	34	35.4	1106	1	US-08-477-329-2	Sequence 2, Appli
703	34	35.4	503	2	US-09-252-991A-17891	Sequence 17891, A	776	34	35.4	1106	1	US-08-475-458-2	Sequence 2, Appli
704	34	35.4	504	2	US-09-071-035-432	Sequence 432, App	777	34	35.4	1106	1	US-08-460-510-2	Sequence 2, Appli
705	34	35.4	504	2	US-10-206-576-432	Sequence 432, App	778	34	35.4	1106	1	US-08-460-490-2	Sequence 2, Appli
706	34	35.4	509	2	US-09-252-991A-20470	Sequence 20470, A	779	34	35.4	1106	2	US-08-980-400-2	Sequence 2, Appli
707	34	35.4	510	2	US-09-211-417-1	Sequence 1, Appli	780	34	35.4	1106	2	US-08-462-728-4	Sequence 4, Appli
708	34	35.4	516	2	US-09-252-991A-21880	Sequence 21880, A	781	34	35.4	1106	2	US-09-583-459A-2	Sequence 2, Appli
709	34	35.4	519	2	US-09-949-016-7883	Sequence 7883, Ap	782	34	35.4	1106	2	US-09-583-210-2	Sequence 2, Appli
710	34	35.4	528	2	US-09-270-767-44826	Sequence 44826, A	783	34	35.4	1106	2	US-09-583-449A-2	Sequence 2, Appli
711	34	35.4	544	2	US-09-902-540-13743	Sequence 13743, A	784	34	35.4	1106	2	US-09-435-059-2	Sequence 2, Appli
712	34	35.4	552	2	US-09-071-035-430	Sequence 430, App	785	34	35.4	1106	2	US-08-461-917-4	Sequence 4, Appli
713	34	35.4	552	2	US-10-206-576-430	Sequence 430, App	786	34	35.4	1106	2	US-08-464-436-4	Sequence 4, Appli
714	34	35.4	561	1	US-08-221-817-16	Sequence 16, Appl	787	34	35.4	1106	2	US-08-464-436-4	Sequence 4, Appli
715	34	35.4	561	1	US-08-454-439-16	Sequence 16, Appl	788	34	35.4	1106	2	US-09-866-510-16	Sequence 16, Appl
716	34	35.4	561	4	PCT-US94-10487-16	Sequence 16, Appl	789	34	35.4	1106	2	US-09-866-510-18	Sequence 18, Appl
717	34	35.4	565	2	US-09-710-279-2608	Sequence 2608, Ap	790	34	35.4	1106	2	US-09-866-510-20	Sequence 20, Appl
718	34	35.4	575	2	US-08-922-865-2	Sequence 2, Appli	791	34	35.4	1106	2	US-09-866-510-22	Sequence 22, Appl
719	34	35.4	575	2	US-09-510-949-2	Sequence 2, Appli	792	34	35.4	1106	4	PCT-US92-00730-2	Sequence 2, Appli
720	34	35.4	588	2	US-09-248-796A-20839	Sequence 20839, A	793	34	35.4	1106	4	PCT-US92-00862-2	Sequence 2, Appli
721	34	35.4	594	2	US-09-489-039A-10369	Sequence 10369, A	794	34	35.4	1109	2	US-09-688-188B-88	Sequence 88, Appl
722	34	35.4	597	2	US-09-252-991A-31010	Sequence 31010, A	795	34	35.4	1109	2	US-09-291-417D-88	Sequence 88, Appl
723	34	35.4	612	2	US-09-489-039A-10140	Sequence 10140, A	796	34	35.4	1128	1	US-08-111-939-2	Sequence 2, Appli
724	34	35.4	617	2	US-09-107-532A-6828	Sequence 6828, Ap	797	34	35.4	1128	2	US-09-641-741-30	Sequence 30, Appl
725	34	35.4	618	2	US-09-248-796A-15319	Sequence 15319, A	798	34	35.4	1128	2	US-09-060-482-8	Sequence 8, Appli
726	34	35.4	622	2	US-09-252-991A-19802	Sequence 19802, A	799	34	35.4	1169	1	US-08-542-921-2	Sequence 2, Appli
727	34	35.4	642	1	US-08-706-936-3	Sequence 3, Appli	800	34	35.4	1169	1	US-08-880-685-2	Sequence 2, Appli
728	34	35.4	645	2	US-09-710-279-1770	Sequence 1770, Ap	801	34	35.4	1169	1	US-08-880-684-2	Sequence 2, Appli
729	34	35.4	648	2	US-09-710-279-568	Sequence 568, App	802	34	35.4	1187	2	US-09-134-001C-5276	Sequence 5276, Ap
730	34	35.4	650	2	US-09-134-001C-3781	Sequence 3781, Ap	803	34	35.4	1187	2	Sequence 6513, Ap	
731	34	35.4	650	2	US-10-104-047-3636	Sequence 3636, Ap	804	34	35.4	1187	2	Sequence 7835, Ap	
732	34	35.4	659	2	US-10-012-819-70	Sequence 70, Appl	805	34	35.4	1747	2	Sequence 7835, Ap	
733	34	35.4	720	2	US-09-252-991A-26650	Sequence 26650, A	806	34	35.4	1806	2	Sequence 56, Appl	
734	34	35.4	723	2	US-09-949-016-10352	Sequence 10352, A	807	34	35.4	1927	2	Sequence 6206, Ap	
735	34	35.4	728	2	US-09-328-352-5642	Sequence 5642, Ap	808	34	35.4	2169	2	Sequence 6930, Ap	
736	34	35.4	736	2	US-09-949-016-7488	Sequence 7488, Ap	809	34	35.4	4861	2	Sequence 6930, Ap	
737	34	35.4	748	2	US-09-538-092-541	Sequence 541, App	810	33.5	34.9	69	2	Sequence 792, App	
738	34	35.4	759	2	US-09-583-110-3807	Sequence 3807, Ap	811	33.5	34.9	69	2	Sequence 792, App	
739	34	35.4	760	2	US-09-107-433-3407	Sequence 3407, Ap	812	33.5	34.9	112	2	Sequence 222, App	
740	34	35.4	773	2	US-09-543-681A-7688	Sequence 7688, Ap	813	33.5	34.9	112	2	Sequence 222, App	
741	34	35.4	795	2	US-09-252-991A-31273	Sequence 31273, A	814	33.5	34.9	112	2	Sequence 222, App	
742	34	35.4	810	2	US-09-198-452A-20	Sequence 20, Appl	815	33.5	34.9	118	2	Sequence 6862, App	
743	34	35.4	810	2	US-09-198-452A-1128	Sequence 1128, Ap	816	33.5	34.9	118	2	Sequence 52179, A	
744	34	35.4	814	2	US-09-252-991A-22923	Sequence 22923, A	817	33.5	34.9	135	2	Sequence 52179, A	
745	34	35.4	819	2	US-09-438-185A-1055	Sequence 1055, Ap	818	33.5	34.9	143	2	Sequence 13936, A	
746	34	35.4	826	2	US-09-134-000C-6598	Sequence 6598, Ap	819	33.5	34.9	151	2	Sequence 2868, Ap	
747	34	35.4	838	2	US-09-270-767-46665	Sequence 46665, A	820	33.5	34.9	188	2	Sequence 551, App	
748	34	35.4	844	2	US-09-252-991A-27184	Sequence 27184, A	821	33.5	34.9	212	2	Sequence 638, App	
749	34	35.4	869	1	US-08-188-582-32	Sequence 32, Appl	822	33.5	34.9	228	2	Sequence 20, Appl	
750	34	35.4	869	1	US-08-646-715-32	Sequence 32, Appl	823	33.5	34.9	228	1	Sequence 19215, A	
751	34	35.4	869	2	US-09-949-016-6649	Sequence 6649, Ap	824	33.5	34.9	249	1	Sequence 57, Appl	
752	34	35.4	875	2	US-09-206-551-14	Sequence 14, Appl	825	33.5	34.9	249	1	Sequence 57, Appl	
753	34	35.4	882	2	US-09-413-844-78	Sequence 78, Appl	826	33.5	34.9	249	1	Sequence 57, Appl	
754	34	35.4	898	1	US-08-465-995A-2	Sequence 2, Appli	827	33.5	34.9	249	2	Sequence 57, Appl	
755	34	35.4	898	1	US-08-465-995A-4	Sequence 4, Appli	828	33.5	34.9	249	2	Sequence 57, Appl	
756	34	35.4	898	1	US-08-465-994C-2	Sequence 2, Appli	829	33.5	34.9	312	1	Sequence 27, Appl	
757	34	35.4	898	1	US-08-465-994C-4	Sequence 4, Appli	830	33.5	34.9	312	1	Sequence 27, Appl	



831	33.5	34.9	312	1	US-08-455-992-27	Sequence 27, Appl	904	33	34.4	32	2	US-10-004-860-643	Sequence 643, Appl
832	33.5	34.9	312	1	US-08-455-972-27	Sequence 27, Appl	905	33	34.4	41	2	US-08-857-076-55	Sequence 55, Appl
833	33.5	34.9	312	4	PCT-US92-00652-27	Sequence 27, Appl	906	33	34.4	41	2	US-09-716-129-112	Sequence 112, Appl
834	33.5	34.9	366	2	US-09-252-991A-16760	Sequence 16760, A	907	33	34.4	41	2	US-09-205-658-55	Sequence 55, Appl
835	33.5	34.9	378	1	US-08-553-367A-6	Sequence 6, Appl	908	33	34.4	41	2	US-09-205-658-168	Sequence 168, Appl
836	33.5	34.9	378	2	US-09-295-306-6	Sequence 6, Appl	909	33	34.4	65	2	US-09-770-595A-18	Sequence 18, Appl
837	33.5	34.9	378	2	US-09-734-719-6	Sequence 7, Appl	910	33	34.4	66	2	US-09-621-976-6086	Sequence 6086, Appl
838	33.5	34.9	378	2	US-10-155-435-7	Sequence 7, Appl	911	33	34.4	70	2	US-09-910-009A-369	Sequence 369, Appl
839	33.5	34.9	397	2	US-09-248-796A-21477	Sequence 21477, A	912	33	34.4	77	2	US-09-621-976-6429	Sequence 6429, Appl
840	33.5	34.9	403	1	US-08-607-509-4	Sequence 4, Appl	913	33	34.4	80	2	US-09-270-767-39725	Sequence 39725, Appl
841	33.5	34.9	403	1	US-08-634-642-4	Sequence 4, Appl	914	33	34.4	80	2	US-09-270-767-54942	Sequence 54942, Appl
842	33.5	34.9	403	2	US-08-989-370-4	Sequence 4, Appl	915	33	34.4	82	1	US-08-497-312-24	Sequence 24, Appl
843	33.5	34.9	403	2	US-09-388-169-4	Sequence 4, Appl	916	33	34.4	83	2	US-08-259-451-9	Sequence 9, Appl
844	33.5	34.9	409	2	US-09-902-540-10657	Sequence 10657, A	917	33	34.4	89	2	US-09-270-767-34138	Sequence 34138, Appl
845	33.5	34.9	410	1	US-08-472-666-2	Sequence 2, Appl	918	33	34.4	89	2	US-09-270-767-49355	Sequence 49355, Appl
846	33.5	34.9	410	2	US-09-101-068-4	Sequence 4, Appl	919	33	34.4	93	2	US-09-248-796A-25152	Sequence 25152, Appl
847	33.5	34.9	410	4	PCT-US96-07615-2	Sequence 2, Appl	920	33	34.4	101	2	US-09-902-540-11494	Sequence 11494, Appl
848	33.5	34.9	446	2	US-09-252-991A-27110	Sequence 27110, A	921	33	34.4	103	2	US-09-270-767-57623	Sequence 57623, Appl
849	33.5	34.9	476	2	US-09-198-452A-1021	Sequence 1021, Appl	922	33	34.4	103	2	US-09-248-796A-17794	Sequence 17794, Appl
850	33.5	34.9	476	2	US-10-197-220-69	Sequence 69, Appl	923	33	34.4	105	2	US-09-252-991A-31703	Sequence 31703, Appl
851	33.5	34.9	479	2	US-09-438-185A-950	Sequence 950, Appl	924	33	34.4	106	2	US-09-252-991A-22105	Sequence 22105, Appl
852	33.5	34.9	502	2	US-09-489-039A-8035	Sequence 8035, Appl	925	33	34.4	107	1	US-07-634-278-34	Sequence 34, Appl
853	33.5	34.9	503	2	US-09-657-013-50	Sequence 50, Appl	926	33	34.4	107	1	US-08-477-728-34	Sequence 34, Appl
854	33.5	34.9	510	2	US-09-248-796A-14490	Sequence 14490, A	927	33	34.4	107	1	US-08-474-040-34	Sequence 34, Appl
855	33.5	34.9	532	2	US-09-949-016-7663	Sequence 7663, A	928	33	34.4	107	1	US-08-487-200-34	Sequence 34, Appl
856	33.5	34.9	549	2	US-09-657-013-54	Sequence 54, Appl	929	33	34.4	107	2	US-08-484-537-34	Sequence 34, Appl
857	33.5	34.9	549	2	US-09-949-016-8881	Sequence 8881, Appl	930	33	34.4	108	2	US-08-752-693A-2	Sequence 2, Appl
858	33.5	34.9	549	2	US-09-949-016-8882	Sequence 8882, Appl	931	33	34.4	109	1	US-07-941-646-21	Sequence 21, Appl
859	33.5	34.9	549	2	US-09-949-016-8883	Sequence 8883, Appl	932	33	34.4	109	1	US-08-147-023-21	Sequence 21, Appl
860	33.5	34.9	549	2	US-09-949-016-8884	Sequence 8884, Appl	933	33	34.4	109	1	US-08-447-570-21	Sequence 21, Appl
861	33.5	34.9	549	2	US-09-949-016-8885	Sequence 8885, Appl	934	33	34.4	109	1	US-08-449-700-21	Sequence 21, Appl
862	33.5	34.9	556	2	US-09-657-013-51	Sequence 51, Appl	935	33	34.4	109	1	US-08-449-700-21	Sequence 21, Appl
863	33.5	34.9	560	1	US-07-683-957B-1	Sequence 1, Appl	936	33	34.4	109	1	US-08-449-699A-21	Sequence 21, Appl
864	33.5	34.9	569	2	US-09-514-245-16	Sequence 16, Appl	937	33	34.4	109	2	US-09-148-925C-21	Sequence 21, Appl
865	33.5	34.9	574	2	US-09-079-431B-6	Sequence 6, Appl	938	33	34.4	109	2	US-08-557-425-21	Sequence 21, Appl
866	33.5	34.9	585	2	US-09-489-039A-14060	Sequence 14060, A	939	33	34.4	109	4	PCT-US93-05446-17	Sequence 17, Appl
867	33.5	34.9	595	2	US-09-949-016-8871	Sequence 8871, Appl	940	33	34.4	109	2	US-09-949-016-9389	Sequence 9389, Appl
868	33.5	34.9	595	2	US-09-949-016-8872	Sequence 8872, Appl	941	33	34.4	111	2	US-09-489-039A-12177	Sequence 12177, Appl
869	33.5	34.9	595	2	US-09-949-016-8873	Sequence 8873, Appl	942	33	34.4	113	2	US-09-270-767-33662	Sequence 33662, Appl
870	33.5	34.9	595	2	US-09-949-016-8874	Sequence 8874, Appl	943	33	34.4	117	2	US-09-270-767-48879	Sequence 48879, Appl
871	33.5	34.9	595	2	US-09-949-016-8875	Sequence 8875, Appl	944	33	34.4	123	2	US-09-640-211A-728	Sequence 728, Appl
872	33.5	34.9	605	2	US-09-657-013-52	Sequence 52, Appl	945	33	34.4	124	2	US-09-252-991A-17879	Sequence 17879, Appl
873	33.5	34.9	629	2	US-09-079-431B-2	Sequence 2, Appl	946	33	34.4	124	2	US-09-252-991A-29476	Sequence 29476, Appl
874	33.5	34.9	630	2	US-09-079-431B-2	Sequence 2, Appl	947	33	34.4	124	2	US-09-270-767-48879	Sequence 48879, Appl
875	33.5	34.9	649	2	US-09-252-991A-31187	Sequence 31187, A	948	33	34.4	126	2	US-08-284-516C-37	Sequence 37, Appl
876	33.5	34.9	651	2	US-09-949-016-8866	Sequence 8866, Appl	949	33	34.4	127	2	US-08-284-516C-41	Sequence 41, Appl
877	33.5	34.9	651	2	US-09-949-016-8867	Sequence 8867, Appl	950	33	34.4	127	2	US-08-284-516C-47	Sequence 47, Appl
878	33.5	34.9	651	2	US-09-949-016-8868	Sequence 8868, Appl	951	33	34.4	127	2	US-08-284-516C-57	Sequence 57, Appl
879	33.5	34.9	651	2	US-09-949-016-8869	Sequence 8869, Appl	952	33	34.4	127	2	US-08-284-516C-65	Sequence 65, Appl
880	33.5	34.9	651	2	US-09-949-016-8870	Sequence 8870, Appl	953	33	34.4	127	2	US-09-537-911A-41	Sequence 41, Appl
881	33.5	34.9	660	2	US-09-949-016-8876	Sequence 8876, Appl	954	33	34.4	127	2	US-09-537-911A-47	Sequence 47, Appl
882	33.5	34.9	660	2	US-09-949-016-8877	Sequence 8877, Appl	955	33	34.4	127	2	US-09-537-911A-57	Sequence 57, Appl
883	33.5	34.9	660	2	US-09-949-016-8878	Sequence 8878, Appl	956	33	34.4	127	2	US-09-537-911A-65	Sequence 65, Appl
884	33.5	34.9	660	2	US-09-949-016-8879	Sequence 8879, Appl	957	33	34.4	128	1	US-08-449-287-6	Sequence 8, Appl
885	33.5	34.9	660	2	US-09-949-016-8880	Sequence 8880, Appl	958	33	34.4	128	1	US-08-449-287-6	Sequence 8, Appl
886	33.5	34.9	820	2	US-09-489-039A-10142	Sequence 10142, A	959	33	34.4	128	1	US-09-270-767-46091	Sequence 46091, Appl
887	33.5	34.9	982	2	US-09-551-974A-95	Sequence 95, Appl	960	33	34.4	128	2	US-09-513-999C-5244	Sequence 5244, Appl
888	33.5	34.9	982	2	US-09-565-501A-95	Sequence 95, Appl	961	33	34.4	133	2	US-09-252-991A-32279	Sequence 32279, Appl
889	33.5	34.9	982	2	US-09-639-206A-95	Sequence 95, Appl	962	33	34.4	136	2	US-08-976-183A-46	Sequence 46, Appl
890	33.5	34.9	982	2	US-09-874-923-95	Sequence 95, Appl	963	33	34.4	138	2	US-09-422-869-14	Sequence 14, Appl
891	33.5	34.9	1641	2	US-09-551-974A-96	Sequence 96, Appl	964	33	34.4	138	2	US-09-902-540-15911	Sequence 15911, Appl
892	33.5	34.9	1641	2	US-09-565-501A-96	Sequence 96, Appl	965	33	34.4	139	2	US-09-902-540-11004	Sequence 11004, Appl
893	33.5	34.9	1641	2	US-09-639-206A-96	Sequence 96, Appl	966	33	34.4	142	2	US-09-252-991A-23667	Sequence 23667, Appl
894	33.5	34.9	1641	2	US-09-874-923-96	Sequence 96, Appl	967	33	34.4	145	2	US-09-252-991A-28281	Sequence 28281, Appl
895	33.5	34.9	2196	2	US-10-360-101-259	Sequence 259, Appl	968	33	34.4	145	2	US-09-252-991A-28281	Sequence 28281, Appl
896	33.5	34.9	2224	2	US-09-054-272-38	Sequence 38, Appl	969	33	34.4	146	1	US-08-688-609-6	Sequence 6, Appl
897	33.5	34.9	2224	2	US-09-949-002-292	Sequence 292, Appl	970	33	34.4	146	2	US-09-002-832-6	Sequence 6, Appl
898	33.5	34.9	15	2	US-08-602-999A-347	Sequence 347, Appl	971	33	34.4	146	2	US-09-976-594-830	Sequence 830, Appl
899	33	34.4	15	2	US-09-500-124-347	Sequence 347, Appl	972	33	34.4	146	2	US-09-788-773A-2	Sequence 2, Appl
900	33	34.4	16	2	US-09-910-009A-448	Sequence 448, Appl	973	33	34.4	146	2	US-09-788-773A-6	Sequence 6, Appl
901	33	34.4	26	1	US-07-942-245-479	Sequence 479, Appl	974	33	34.4	146	2	US-09-788-773A-8	Sequence 8, Appl
902	33	34.4	29	2	US-09-270-767-59495	Sequence 59495, A	975	33	34.4	146	2	US-09-788-773A-8	Sequence 8, Appl
903	33	34.4	32	2	US-09-205-258-643	Sequence 643, Appl	976	33	34.4	148	1	US-08-688-609-3	Sequence 3, Appl

977 33 34.4 148 2 US-09-002-832-3  
978 33 34.4 154 2 US-09-252-991A-31454  
979 33 34.4 155 2 US-09-252-991A-23857  
980 33 34.4 159 2 US-08-390-353A-5  
981 33 34.4 159 2 US-09-101-059-5  
982 33 34.4 160 2 US-09-270-767-42334  
983 33 34.4 166 2 US-09-917-254-99  
984 33 34.4 167 2 US-09-170-984-14  
985 33 34.4 169 2 US-08-483-533-28  
986 33 34.4 169 2 US-09-283-471A-28  
987 33 34.4 176 2 US-09-248-796A-18992  
988 33 34.4 176 2 US-09-489-039A-7497  
989 33 34.4 177 2 US-09-248-796A-21436  
990 33 34.4 181 2 US-08-529-055-42  
991 33 34.4 182 2 US-09-540-236-2367  
992 33 34.4 183 2 US-08-868-452-38  
993 33 34.4 184 2 US-09-605-703B-1864  
994 33 34.4 184 2 US-09-938-391A-11  
995 33 34.4 188 2 US-09-252-991A-28878  
996 33 34.4 189 1 US-08-026-758-16  
997 33 34.4 189 2 US-09-270-767-44080  
998 33 34.4 196 2 US-09-949-016-8932  
999 33 34.4 196 2 US-09-949-016-8933  
1000 33 34.4 196 2 US-09-949-016-8934

## ALIGNMENTS

RESULT 1  
US-08-208-008C-13  
; Sequence 13, Application US/08208008C  
; Patent No. 5633147  
; GENERAL INFORMATION:  
; APPLICANT: MEISSNER, ET AL.  
; TITLE OF INVENTION: Transforming Growth Factor  
; TITLE OF INVENTION: Alpha - H1  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/208.008C  
; FILING DATE: March 8, 1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MULLINS, J.G.  
; REGISTRATION NUMBER: 33,073  
; REFERENCE/DOCKET NUMBER: 325800-98  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 68 AMINO ACIDS  
; TYPE: AMINO ACID  
; STRANDEDNESS:  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; US-08-208-008C-13

Query Match 100.0%; Score 96; DB 1; Length 68;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 FRDDSIWQPEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | |  
Db 20 FRDDSIWQPEPAIRPR 36  
| | | | | | | | | | | | | | | | | | | | |  
  
RESULT 2  
5256643-3  
; Patent No. 5256643  
; APPLICANT: Persico, Maria G.; Salomon, David S.  
; TITLE OF INVENTION: HUMAN CRYPTO PROTEIN  
; NUMBER OF SEQUENCES: 18  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/530.165  
; FILING DATE: 29-MAY-1990  
; SEQ ID NO: 3  
; LENGTH: 174  
5256643-3  
  
Query Match 100.0%; Score 96; DB 6; Length 174;  
Best Local Similarity 100.0%; Pred. No. 5.8e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 FRDDSIWQPEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | |  
Db 32 FRDDSIWQPEPAIRPR 48  
| | | | | | | | | | | | | | | | | | | | |

RESULT 3  
5256643-2  
; Patent No. 5256643  
; APPLICANT: Persico, Maria G.; Salomon, David S.  
; TITLE OF INVENTION: HUMAN CRYPTO PROTEIN  
; NUMBER OF SEQUENCES: 18  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/530.165  
; FILING DATE: 29-MAY-1990  
; SEQ ID NO: 2  
; LENGTH: 187  
5256643-2  
  
Query Match 100.0%; Score 96; DB 6; Length 187;  
Best Local Similarity 100.0%; Pred. No. 6.2e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 FRDDSIWQPEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | |  
Db 46 FRDDSIWQPEPAIRPR 62  
| | | | | | | | | | | | | | | | | | | | |

RESULT 4  
US-07-749-001-3  
; Sequence 3, Application US/07749001  
; Patent No. 5264557  
; GENERAL INFORMATION:  
; APPLICANT: Salomon, David S.  
; APPLICANT: Persico, Maria G.  
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICANT: SALOMON, David S.  
FILING DATE: 19910823  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Scott, Watson T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 248453 CUSH  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-749-001-3

Query Match 100.0%; Score 96; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6.3e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 5  
US-07-749-001-5  
Sequence 5, Application US/07749001  
Patent No. 5264557  
GENERAL INFORMATION:  
APPLICANT: SALOMON, David S.  
APPLICANT: PERSICO, Maria G.  
TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARRY & CUSHMAN  
STREET: 1615 L Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICANT: SALOMON, David S.  
FILING DATE: 19910823  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Scott, Watson T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 248453 CUSH  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-749-001-5

Query Match 100.0%; Score 96; DB 1; Length 188;

Best Local Similarity 100.0%; Pred. No. 6.3e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 FRDSDIWPQEEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 46 FRDSDIWPQEEPAIRPR 62  
RESULT 6  
US-08-154-198-3  
Sequence 3, Application US/08154198  
Patent No. 5620866  
GENERAL INFORMATION:  
APPLICANT: SALOMON, David S.  
APPLICANT: PERSICO, Maria G.  
TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Steuart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICANT: SALOMON, David S.  
FILING DATE: 17-NOV-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,001  
FILING DATE: 23-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 15280-63-1  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-154-198-3

Query Match 100.0%; Score 96; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6.3e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 7  
US-08-154-198-5  
Sequence 5, Application US/08154198  
Patent No. 5620866  
GENERAL INFORMATION:  
APPLICANT: SALOMON, David S.  
APPLICANT: PERSICO, Maria G.  
TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Steuart Street Tower, One Market Plaza

CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/154,198  
FILING DATE: 17-NOV-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,001  
FILING DATE: 23-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 15280-63-1  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-154-198-5

Query Match 100.0%; Score 96; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6.3e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSDIWQEEPAIRPR 17  
|||||  
DB 46 FRDSDIWQEEPAIRPR 62

RESULT 8  
US-08-463-335-3  
Sequence 3, Application US/08463335  
Patent No. 5650285  
GENERAL INFORMATION:  
APPLICANT: Salomon, David S.  
APPLICANT: Persico, Maria G.  
TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARRY & CUSHMAN  
STREET: 1615 L Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,335  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,001  
FILING DATE: 23-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Scott, Watson T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: 15280-63-1  
TELEPHONE: (202) 822-0944  
TELEFAX: (202) 822-0944  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-463-335-5

TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 248453 CUSH  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-463-335-3

Query Match 100.0%; Score 96; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6.3e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSDIWQEEPAIRPR 17  
|||||  
DB 46 FRDSDIWQEEPAIRPR 62

RESULT 9  
US-08-463-335-5  
Sequence 5, Application US/08463335  
Patent No. 5650285  
GENERAL INFORMATION:  
APPLICANT: Salomon, David S.  
APPLICANT: Persico, Maria G.  
TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARRY & CUSHMAN  
STREET: 1615 L Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,335  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,001  
FILING DATE: 23-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Scott, Watson T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 248453 CUSH  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-463-335-5

Query Match 100.0%; Score 96; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6.3e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSDIWQEEPAIRPR 17  
|||||  
DB 46 FRDSDIWQEEPAIRPR 62

RESULT 10  
US-08-464-023A-3  
; Sequence 3, Application US/08464023A  
; Patent No. 5854399  
; GENERAL INFORMATION:  
; APPLICANT: Salomon, David S.  
; APPLICANT: Persico, Maria G.  
; TITLE OF INVENTION: A Human Crypto-Related Gene  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew  
; STREET: One Market Plaza, Steuart Street Tower  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105-1492  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/464,023A  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,001  
; FILING DATE: 23-AUG-1995  
; APPLICATION NUMBER: 15280-63-1-1  
; FILING DATE: 17-NOV-1995  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 15280-63-1-1  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 188 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-464-023A-3  
Query Match 100.0%; Score 96; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6.3e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FRDTSIWPQEEPAIRPR 17  
Db 46 FRDTSIWPQEEPAIRPR 62  
RESULT 11  
US-08-464-023A-5  
; Sequence 5, Application US/08464023A  
; Patent No. 5854399  
; GENERAL INFORMATION:  
; APPLICANT: Salomon, David S.  
; APPLICANT: Persico, Maria G.  
; TITLE OF INVENTION: A Human Crypto-Related Gene  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew  
; STREET: One Market Plaza, Steuart Street Tower  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105-1492  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471,371  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ferraro, Gregory D.  
; REGISTRATION NUMBER: 36,134  
; REFERENCE/DOCKET NUMBER: 32580-455  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (201) 994-1700  
; TELEFAX: (201) 994-1744  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 188 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-464-023A-3

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,023A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,001  
FILING DATE: 23-AUG-1995  
APPLICATION NUMBER: 15280-63-1-1  
FILING DATE: 17-NOV-1995  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 15280-63-1-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-464-023A-5  
Query Match 100.0%; Score 96; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6.3e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FRDTSIWPQEEPAIRPR 17  
Db 46 FRDTSIWPQEEPAIRPR 62  
RESULT 12  
US-08-471-371-7  
; Sequence 7, Application US/08471371  
; Patent No. 5981215  
; GENERAL INFORMATION:  
; APPLICANT: Meissner, Paul S.  
; APPLICANT: Coleman, Timothy A.  
; TITLE OF INVENTION: HUMAN CRIPTIN GROWTH FACTOR  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
; ADDRESSEE: STEWART & OLSTEIN  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: N.J.  
; COUNTRY: U.S.A  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471,371  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ferraro, Gregory D.  
; REGISTRATION NUMBER: 36,134  
; REFERENCE/DOCKET NUMBER: 32580-455  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (201) 994-1700  
; TELEFAX: (201) 994-1744  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 188 amino acids

;  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-471-371-7

Query Match 100.0%; Score 96; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6.3e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWQEEPAIRPR 17  
Db 46 FRDSIWQEEPAIRPR 62  
|||||

## RESULT 13

US-09-949-016-7062  
; Sequence 7062, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7062  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-7062

Query Match 100.0%; Score 96; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 6.4e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWQEEPAIRPR 17  
Db 50 FRDSIWQEEPAIRPR 66  
|||||

## RESULT 14

US-09-949-016-8959  
; Sequence 8959, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8959  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-8959

Query Match 100.0%; Score 96; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 6.4e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWQEEPAIRPR 17  
Db 50 FRDSIWQEEPAIRPR 66  
|||||

## RESULT 15

US-09-949-016-11113  
; Sequence 11113, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11113  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11113

Query Match 100.0%; Score 96; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 6.4e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWQEEPAIRPR 17  
Db 50 FRDSIWQEEPAIRPR 66  
|||||

## RESULT 16

US-08-441-629-7  
; Sequence 7, Application US/08441629  
; Patent No. 5766923  
; GENERAL INFORMATION:  
; APPLICANT: Kirschner, Marc W.  
; APPLICANT: Kinoshita, No. 5766923iyuki  
; TITLE OF INVENTION: RECEPTOR-LIGAND ASSAY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/441,629  
; FILING DATE: 15-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; PRIOR APPLICATION NUMBER: US 08/279,217  
; FILING DATE: 22-JUL-1994  
; ATTORNEY/AGENT INFORMATION:

```
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: HU95-01A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 160 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-441-629-7

Query Match 66.7%; Score 64; DB 1; Length 160;
Best Local Similarity 68.8%; Pred. No. 0.032;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RDSIWQEEPAIRPR 17
Db 31 RDNSWDQKEPAVRDR 46

RESULT 17
US-08-776-207-7
; Sequence 7, Application US/08776207A
; Patent No. 6080718
; GENERAL INFORMATION:
; APPLICANT: Kirschner, Marc W.
; APPLICANT: Kinoshita, No. 6080718iyuki
; TITLE OF INVENTION: Receptor-Ligand Assay
; FILE REFERENCE: HU95-01A2
; CURRENT APPLICATION NUMBER: US/08/776.207A
; CURRENT FILING DATE: 1997-06-23
; EARLIER APPLICATION NUMBER: PCT/US95/09172
; EARLIER FILING DATE: 1995-07-19
; EARLIER APPLICATION NUMBER: 08/441.629
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/279,217
; EARLIER FILING DATE: 1994-07-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-776-207-7

Query Match 66.7%; Score 64; DB 2; Length 160;
Best Local Similarity 68.8%; Pred. No. 0.032;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RDSIWQEEPAIRPR 17
Db 31 RDNSWDQKEPAVRDR 46

RESULT 18
US-09-507-773-7
; Sequence 7, Application US/09507773
; Patent No. 6399386
; GENERAL INFORMATION:
; APPLICANT: Kirschner, Marc W.
; APPLICANT: Kinoshita, No. 6399386iyuki
; TITLE OF INVENTION: Receptor-Ligand Assay
; FILE REFERENCE: HU95-01A2
; CURRENT APPLICATION NUMBER: US/09/507,773
; CURRENT FILING DATE: 2000-02-18
; EARLIER APPLICATION NUMBER: 08/776,207
; PRIOR FILING DATE: 1997-06-23
; PRIOR APPLICATION NUMBER: 08/441,629
; PRIOR FILING DATE: 1995-05-15
```

```
; PRIOR APPLICATION NUMBER: 08/279,217
; PRIOR FILING DATE: 1994-07-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-507-773-7

Query Match 66.7%; Score 64; DB 2; Length 160;
Best Local Similarity 68.8%; Pred. No. 0.032;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RDSIWQEEPAIRPR 17
Db 31 RDNSWDQKEPAVRDR 46

RESULT 19
US-10-016-447-7
; Sequence 7, Application US/10016447
; Patent No. 6844193
; GENERAL INFORMATION:
; APPLICANT: Kirschner, Marc W.
; APPLICANT: Kinoshita, No. 6844193iyuki
; TITLE OF INVENTION: Receptor-Ligand Assay
; FILE REFERENCE: HU95-01A2
; CURRENT APPLICATION NUMBER: US/10/016,447
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/08/776,207
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-06-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/441,629
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-05-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/279,217
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-07-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-016-447-7

Query Match 66.7%; Score 64; DB 2; Length 160;
Best Local Similarity 68.8%; Pred. No. 0.032;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RDSIWQEEPAIRPR 17
Db 31 RDNSWDQKEPAVRDR 46

RESULT 20
PCT-US95-09172-7
; Sequence 7, Application PC/TUS9509172
; GENERAL INFORMATION:
; APPLICANT: Kirschner, Marc W.
; APPLICANT: Kinoshita, Noriyuki
; TITLE OF INVENTION: RECEPTOR-LIGAND ASSAY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09172
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/279,217
FILING DATE: 23-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/441,629
FILING DATE: 15-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: H095-01A PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-09172-7

```

Query Match 66.7%; Score 64; DB 4; Length 160;  
Best Local Similarity 68.8%; Pred. No. 0.032;  
Matches 11; Conservative 3; Mismatches 2; Indels

Qy 2 R D S I W P Q E E P A I R P R 17  
||:|||||:||||:|  
Db 31 R D N S I W D Q K E P A V R D R 46

```

RESULT 21
US-09-252-991A-18373
/ Sequence 18373, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND
/ TITLE OF INVENTION: AERUGINOSA FOR DIAG
/ FILE REFERENCE: 107196:136
/ CURRENT APPLICATION NUMBER: US/09/252,991
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,199
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 18373
/ LENGTH: 252
/ TYPE: PRP
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18373

```

Query Match 47.9%; Score 46; DB 2; Length 252;  
Best Local Similarity 60.0%; Pred. No. 25;  
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 R D D S I W P Q E E P A I R P 16  
D<sub>b</sub> 105 R G R S P W P P A E P A I H P 119

RESULT 22  
US-09-270-767-43850  
; Sequence 43850, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and

```

; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US-09-270-767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43850
; LENGTH: 386
; TYPE: PRN
; ORGANISM: Drosophila melanogaster
US-09-270-767-43850

```

Query Match 46.9%; Score 45; DB 2; Length 386;  
Best Local Similarity 50.0%; Pred. No. 55;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy	1	FRDDSIWPQEEP	14
		:   :	
pb	159	FOEDDVWSOEDE	172

RESULT 23  
US-09-252-991A-30230  
; Sequence 30230, Application US/09252991A  
; Patent NO. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30230  
; LENGTH: 448  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30230

Query Match	46.9%;	Score 45;	DB 2;	Length 448;
Best Local Similarity	77.8%;	Pred. No. 65;		
Matches 7; Conservative	1;	Mismatches 1;	Indels 0;	Gaps 0;

Qy 7 WPQEEPAIR 15  
Db 182 WPGEPAVR 190

```

RESULT 24
US-09-538-092-216
; Sequence 216, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 216
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:

```



```
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YDR490C
US-09-538-092-216

Query Match      46.9%; Score 45; DB 2; Length 766;
Best Local Similarity 43.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 FRDSIWPOEPAIRP 16
Db 396 FEDGSVDDNPPEIQP 411

RESULT 25
US-09-248-796A-19171
; Sequence 19171, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19171
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19171

Query Match      45.8%; Score 44; DB 2; Length 240;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 RDSIWPOEPPA 13
Db 209 QDESPLSEPS 220

RESULT 26
US-09-894-998A-51
; Sequence 51, Application US/09894998A
; Patent No. 6537555
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Dav
; APPLICANT: Davin C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 210121.538
; CURRENT APPLICATION NUMBER: US/09/894,998A
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 1113
; TYPE: PRT
; ORGANISM: HSV-2
US-09-894-998A-51

Query Match      44.8%; Score 43; DB 2; Length 1113;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 DDSIWPOEPAIRP 16
```

```
Db 1100 DDEMTPOETPPVHP 1113

RESULT 27
US-10-237-551-51
; Sequence 51, Application US/10237551
; Patent No. 6821519
; GENERAL INFORMATION:
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 210121.538C3
; CURRENT APPLICATION NUMBER: US/10/237,551
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 1113
; TYPE: PRT
; ORGANISM: HSV-2
US-10-237-551-51

Query Match      44.8%; Score 43; DB 2; Length 1113;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 DDSIWPOEPAIRP 16
Db 1100 DDEMTPOETPPVHP 1113

RESULT 28
US-10-237-551-202
; Sequence 202, Application US/10237551
; Patent No. 6821519
; GENERAL INFORMATION:
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 210121.538C3
; CURRENT APPLICATION NUMBER: US/10/237,551
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202
; LENGTH: 1114
; TYPE: PRT
; ORGANISM: HSV2
US-10-237-551-202

Query Match      44.8%; Score 43; DB 2; Length 1114;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 DDSIWPOEPAIRP 16
Db 1101 DDEMTPOETPPVHP 1114

RESULT 29
US-09-489-039A-10445
; Sequence 10445, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
```

; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 10445  
; LENGTH: 255  
; TYPE: PRF  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-10445

Query Match 43.8%; Score 42; DB 2; Length 255;  
Best Local Similarity 54.5%; Pred. No. 1e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 7 WQEEPAIRPR 17  
||:|:|:  
Db 174 WPADNPDIKPR 184

## RESULT 30

US-09-252-991A-29455  
; Sequence 29455, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 29455  
; LENGTH: 349  
; TYPE: PRF  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29455

Query Match 43.8%; Score 42; DB 2; Length 349;  
Best Local Similarity 58.3%; Pred. No. 1.4e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 6 IWQEEPAIRPR 17  
:|:|:|:  
Db 53 LFPQSPATPR 64

## RESULT 31

US-08-375-134-12  
; Sequence 12, Application US/08375134  
; Patent No. 5602011  
; GENERAL INFORMATION:  
; APPLICANT: Luhm, Robert A.  
; APPLICANT: d'Hennezel, Olga B.  
; APPLICANT: Duffaud, Guy D.  
; APPLICANT: Jolly, James F.  
; APPLICANT: Kelly, Robert M.  
; APPLICANT: Ting, Eve Yi-Fay  
; TITLE OF INVENTION: PURIFIED THERMOCOCCUS BAROSSII DNAPOLYMERASE  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Quarles & Brady  
; STREET: 411 East Wisconsin Avenue  
; CITY: Milwaukee  
; STATE: Wisconsin  
; COUNTRY: U.S.A.  
; ZIP: 53202-4497  
; COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/375,134  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baker, Jean C.  
; REGISTRATION NUMBER: 35,433  
; REFERENCE/DOCKET NUMBER: 740211.90440  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (414) 277-5709  
; TELEFAX: (414) 271-3552  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 779 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-375-134-12

Query Match 43.8%; Score 42; DB 1; Length 779;  
Best Local Similarity 42.9%; Pred. No. 3.2e+02;  
Matches 9; Conservative 2; Mismatches 4; Indels 6; Gaps 1;

Qy 2 RDSIIW-----PQEEPAIRP 16  
|:|:  
Db 78 RSVEVWVLYFTHPDVPAIRP 98

## RESULT 32

PCT-US95-15263-12  
; Sequence 12, Application PC/TUS9515263  
; GENERAL INFORMATION:  
; APPLICANT: Luhm, Robert A.  
; APPLICANT: d'Hennezel, Olga B.  
; APPLICANT: Duffaud, Guy D.  
; APPLICANT: Jolly, James F.  
; APPLICANT: Kelly, Robert M.  
; APPLICANT: Ting, Eve Yi-Fay  
; TITLE OF INVENTION: PURIFIED THERMOCOCCUS BAROSSII DNAPOLYMERASE  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Quarles & Brady  
; STREET: 411 East Wisconsin Avenue  
; CITY: Milwaukee  
; STATE: Wisconsin  
; COUNTRY: U.S.A.  
; ZIP: 53202-4497  
; COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/15263  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baker, Jean C.  
; REGISTRATION NUMBER: 35,433  
; REFERENCE/DOCKET NUMBER: 740211.90440  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (414) 277-5709  
; TELEFAX: (414) 271-3552  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 779 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
PCT-US95-15263-12

Query Match 43.8%; Score 42; DB 4; Length 779;  
Best Local Similarity 42.9%; Pred. No. 3.2e+02;  
Matches 9; Conservative 2; Mismatches 4; Indels 6; Gaps 1;

Qy 2 RDSIW-----PQEEPAIRP 16  
| : : : : :  
Db 78 RSEVWVLYFTHPQDVPAIRP 98

## RESULT 33

US-09-919-497-98  
; Sequence 98, Application US/09919497  
; Patent No. 6773883  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George L.  
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER  
; FILE REFERENCE: B0801/7225  
; CURRENT APPLICATION NUMBER: US/09/919,497  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/221,735  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 98  
; LENGTH: 1736  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-919-497-98

Query Match 43.8%; Score 42; DB 2; Length 1736;  
Best Local Similarity 43.8%; Pred. No. 7.4e+02;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 RDSIWQEEPAIRP 17  
| : : : : :  
Db 138 RSEKIWRDRSASR 153

## RESULT 34

US-09-949-016-9470  
; Sequence 9470, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9470  
; LENGTH: 1771  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9470

Query Match 43.8%; Score 42; DB 2; Length 1771;  
Best Local Similarity 43.8%; Pred. No. 7.5e+02;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 RDSIWQEEPAIRP 17  
| : : : : :  
Db 173 RSEKIWRDRSASR 188

## RESULT 35

US-09-248-796A-17216  
; Sequence 17216, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 17216  
; LENGTH: 226  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-17216

Query Match 42.7%; Score 41; DB 2; Length 226;  
Best Local Similarity 60.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 DSIWQEEPAIRP 12  
| : : : : :  
Db 156 EDLWVTOEPP 165

## RESULT 36

US-09-489-039A-12982  
; Sequence 12982, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 12982  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-12982

Query Match 42.7%; Score 41; DB 2; Length 590;  
Best Local Similarity 60.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 DSIWQEEPA 13  
| : : : : :  
Db 27 DDLWPQARPA 36

## RESULT 37

US-08-895-522-3  
; Sequence 3, Application US/08895522  
; Patent No. 5858719  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Shah, Purvi  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE  
; TITLE OF INVENTION: TRANSPORT PROTEIN  
; NUMBER OF SEQUENCES: 4

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895,522
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0336 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 694 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1167982
US-08-895-522-3

Query Match 42.7%; Score 41; DB 1; Length 694;
Best Local Similarity 41.7%; Pred. No. 4e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 6 IWPOEPAIRPR 17
Db 70 WVPEDPDLRAR 81

RESULT 38
US-09-195-391-3
Sequence 3, Application US/09195391
Patent No. 6080842
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Shah, Purvi
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE
TITLE OF INVENTION: TRANSPORT PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/195,391
FILING DATE:

```

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/895,522
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0336 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 694 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1167982
US-09-195-391-3

Query Match 42.7%; Score 41; DB 2; Length 694;
Best Local Similarity 41.7%; Pred. No. 4e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 6 IWPOEPAIRPR 17
Db 70 WVPEDPDLRAR 81

RESULT 39
US-09-240-274-29
Sequence 29, Application US/09240274
Patent No. 6255455
GENERAL INFORMATION:
APPLICANT: Siegel, Donald L.
TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF
FILE REFERENCE: 09596-4202
CURRENT APPLICATION NUMBER: US/09/240,274
CURRENT FILING DATE: 1999-01-29
EARLIER APPLICATION NUMBER: 60/081,380
EARLIER FILING DATE: 1998-04-10
EARLIER APPLICATION NUMBER: 60/028,550
EARLIER FILING DATE: 1996-10-11
NUMBER OF SEQ ID NOS: 224
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 29
LENGTH: 107
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: anti-Rh(D) chain F01
US-09-240-274-29

Query Match 41.7%; Score 40; DB 2; Length 107;
Best Local Similarity 41.2%; Pred. No. 82;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 FRDSDIWPOEPAIRPR 17
Db 28 FRNDLGHVQKPGKAPK 44

RESULT 40
US-09-848-798-29
Sequence 29, Application US/09848798
Patent No. 6858719
GENERAL INFORMATION:
APPLICANT: Siegel, Donald L.
TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL

```

; TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF  
; FILE REFERENCE: 09596-4202  
; CURRENT APPLICATION NUMBER: US/09/848,798  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/240,274  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/028,550  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-10-11  
; NUMBER OF SEQ ID NOS: 224  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 29  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: anti-Rh(D) chain F01  
US-09-848-798-29

Query Match 41.7%; Score 40; DB 2; Length 107;  
Best Local Similarity 41.2%; Pred. No. 82;  
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRDLSIWPOEPAIRPR 17  
||:| |::| |:  
Db 28 FRNDLGWYQKPKAKP 44

RESULT 41  
US-09-726-219A-244  
; Sequence 244, Application US/09726219A  
; Patent No. 6806079  
; GENERAL INFORMATION:  
; APPLICANT: Cambridge Antibody Technology  
; APPLICANT: Medical Research Council  
; APPLICANT: McCafferty, John  
; APPLICANT: Pope, Anthony  
; APPLICANT: Johnson, Kevin  
; APPLICANT: Hoogenboom, Hendricus  
; APPLICANT: Griffiths, Andrew  
; APPLICANT: Jackson, Ronald  
; APPLICANT: Holliger, Kasper  
; APPLICANT: Marks, James  
; APPLICANT: Clackson, Timothy  
; APPLICANT: Chiswell, David  
; APPLICANT: Winter, Gregory  
; APPLICANT: Bonert, Timothy  
; TITLE OF INVENTION: Methods for Producing Members of Specific Binding Pairs  
; FILE REFERENCE: 213839-00013  
; CURRENT APPLICATION NUMBER: US/09/726,219A  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: GB 9015198.6  
; PRIOR FILING DATE: 1990-07-10  
; PRIOR APPLICATION NUMBER: GB 9022845.3  
; PRIOR FILING DATE: 1990-10-19  
; PRIOR APPLICATION NUMBER: GB 9022845.3  
; PRIOR FILING DATE: 1990-10-19  
; PRIOR APPLICATION NUMBER: GB 9024503.6  
; PRIOR FILING DATE: 1990-11-12  
; PRIOR APPLICATION NUMBER: GB 9104744.9  
; PRIOR FILING DATE: 1991-03-06  
; PRIOR APPLICATION NUMBER: GB 9110549.4  
; PRIOR FILING DATE: 1991-05-15  
; PRIOR APPLICATION NUMBER: PCT/GB91/01134  
; PRIOR FILING DATE: 1991-07-10  
; PRIOR APPLICATION NUMBER: US 07/971,857  
; PRIOR FILING DATE: 1993-01-08  
; PRIOR APPLICATION NUMBER: US 08/484,893  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 272  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 244  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: VL of scFv from mouse immunized with 2-phenyl-5-oxazolone  
US-09-848-798-29

; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: VL of scFv from mouse immunized with 2-phenyl-5-oxazolone  
US-09-726-219A-244

Query Match 41.7%; Score 40; DB 2; Length 108;  
Best Local Similarity 45.5%; Pred. No. 83;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 WPOEPAIRPR 17  
||:| |::| |:  
Db 34 WPOQKEGTSFK 44

RESULT 42  
US-09-196-522-244  
; Sequence 244, Application US/09196522  
; Patent No. 6916605  
; GENERAL INFORMATION:  
; APPLICANT: Cambridge Antibody Technology  
; APPLICANT: Medical Research Council  
; APPLICANT: McCafferty, John  
; APPLICANT: Pope, Anthony  
; APPLICANT: Johnson, Kevin  
; APPLICANT: Hoogenboom, Hendricus  
; APPLICANT: Griffiths, Andrew  
; APPLICANT: Jackson, Ronald  
; APPLICANT: Holliger, Kasper  
; APPLICANT: Marks, James  
; APPLICANT: Clackson, Timothy  
; APPLICANT: Chiswell, David  
; APPLICANT: Winter, Gregory  
; APPLICANT: Bonert, Timothy  
; TITLE OF INVENTION: Methods for Producing Members of Specific Binding Pairs  
; FILE REFERENCE: 213839-00004  
; CURRENT APPLICATION NUMBER: US/09/196,522  
; CURRENT FILING DATE: 1998-11-28  
; PRIOR APPLICATION NUMBER: GB 9015198.6  
; PRIOR FILING DATE: 1990-07-10  
; PRIOR APPLICATION NUMBER: GB 9022845.3  
; PRIOR FILING DATE: 1990-10-19  
; PRIOR APPLICATION NUMBER: GB 9022845.3  
; PRIOR FILING DATE: 1990-10-19  
; PRIOR APPLICATION NUMBER: GB 9024503.6  
; PRIOR FILING DATE: 1990-11-12  
; PRIOR APPLICATION NUMBER: GB 9104744.9  
; PRIOR FILING DATE: 1991-03-06  
; PRIOR APPLICATION NUMBER: GB 9110549.4  
; PRIOR FILING DATE: 1991-05-15  
; PRIOR APPLICATION NUMBER: PCT/GB91/01134  
; PRIOR FILING DATE: 1991-07-10  
; PRIOR APPLICATION NUMBER: US 07/971,857  
; PRIOR FILING DATE: 1993-01-08  
; PRIOR APPLICATION NUMBER: US 08/484,893  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 272  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 244  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: VL of scFv from mouse immunized with 2-phenyl-5-oxazolone  
US-09-196-522-244

Query Match 41.7%; Score 40; DB 2; Length 108;  
Best Local Similarity 45.5%; Pred. No. 83;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 WPOEPAIRPR 17  
||:| |::| |:  
Db 34 WPOQKEGTSFK 44

Db 34 WPOQKPGTSPK 44

## RESULT 43

US-09-107-532A-5453  
; Sequence 5453, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 5453:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 126 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...126  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5453:

US-09-107-532A-5453  
Query Match 41.7%; Score 40; DB 2; Length 126;  
Best Local Similarity 42.9%; Pred. No. 97;  
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
Qy 2 RDSIWPOQEEPAIR 15  
Db 44 RHDELWPEMRTALR 57

## RESULT 44

US-09-489-039A-8476  
; Sequence 8476, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 8476  
; LENGTH: 154  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-8476

Query Match 41.7%; Score 40; DB 2; Length 154;  
Best Local Similarity 40.0%; Pred. No. 1.2e+02;  
Matches 8; Conservative 3; Mismatches 5; Indels 4; Gaps 1;

Qy 1 FRDSI---WPOEPAIRP 16  
Db 100 FADDAICLVEWPOQGAGVLP 119

## RESULT 45

US-09-640-211A-834  
; Sequence 834, Application US/09640211A  
; Patent No. 6833446  
; GENERAL INFORMATION:

; APPLICANT: Wood, Marion  
; APPLICANT: Shenk, Michael A.  
; APPLICANT: McGrath, Annette  
; APPLICANT: Glenn, Matthew  
; TITLE OF INVENTION: Compositions and Methods for the  
; TITLE OF INVENTION: Modification of Gene Transcription  
; FILE REFERENCE: 11000.1021CIU  
; CURRENT APPLICATION NUMBER: US/09/640,211A  
; CURRENT FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 2368  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 834  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Eucalyptus grandis  
US-09-640-211A-834

Query Match 41.7%; Score 40; DB 2; Length 180;  
Best Local Similarity 38.5%; Pred. No. 1.4e+02;  
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DSIWPOEPAIRP 16  
Db 37 DSVFPKQPNLNP 49

## RESULT 46

US-09-252-991A-16962  
; Sequence 16962, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 16962  
; LENGTH: 186  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16962

Query Match 41.7%; Score 40; DB 2; Length 186;

Best Local Similarity 61.5%; Pred. No. 1.5e+02; 2; Indels 1;  
Matches 8; Conservative 1; Mismatches 2; Gaps 1;

QY 7 WPOEPAI--RPR 17  
Db 12 WPCPEPAVGRPR 24

## RESULT 47

US-10-104-047-2017  
; Sequence 2017, Application US/10104047  
; Patent No. 6943241  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. 6943241el full length cDNA  
; FILE REFERENCE: HI-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; CURRENT FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2017  
; LENGTH: 263  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-104-047-2017

Query Match 41.7%; Score 40; DB 2; Length 263;  
Best Local Similarity 54.5%; Pred. No. 2.1e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 WPOEPAI--RPR 17  
Db 207 WPCDPPQDRPX 217

## RESULT 48

US-09-252-991A-20824  
; Sequence 20824, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20824  
; LENGTH: 418  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20824

Query Match 41.7%; Score 40; DB 2; Length 418;  
Best Local Similarity 53.8%; Pred. No. 3.4e+02;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 RDDSINWQEEPAI 14  
Db 287 REISTWPTQPAI 299

## RESULT 49

US-09-902-540-14431  
; Sequence 14431, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 14431  
; LENGTH: 471  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-14431

Query Match 41.7%; Score 40; DB 2; Length 471;  
Best Local Similarity 63.8%; Pred. No. 3.8e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 WPOEPAI--RPR 17  
Db 413 WPCGPGIRAR 423

## RESULT 50

US-09-902-540-14372  
; Sequence 14372, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 14372  
; LENGTH: 614  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-14372

Query Match 41.7%; Score 40; DB 2; Length 614;  
Best Local Similarity 53.8%; Pred. No. 5e+02;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 DDSINWQEEPAI 15  
Db 220 DDELWRTKPAIR 232

Search completed: February 3, 2006, 18:57:58  
Job time : 43.5 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:39:57 ; Search time 152.5 Seconds  
(without alignments)  
78.649 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_46\_62

Perfect score: 96

Sequence: 1 FRDSIWPOEFAIRPR 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Uniprot\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	188	1	TDGF1_HUMAN
2	96	100.0	188	1	TDGF2_HUMAN
3	96	100.0	188	2	Q8TCC1_HUMAN
4	70	72.9	181	2	Q5BD57_BOVIN
5	64	66.7	171	1	TDGF1_MOUSE
6	64	66.7	171	2	Q7TQ06_MOUSE
7	49	51.0	241	2	Q96ZC2_SULTO
8	49	51.0	1312	2	Q4WQF8_ASPTU
9	49	51.0	1319	2	Q5B761_EMENI
10	49	51.0	1553	2	Q4JQ52_CHLRE
11	48	50.0	333	2	Q60JD3_CAEBR
12	48	50.0	333	2	Q22352_CAEBL
13	48	50.0	709	2	Q7KVA9_DROME
14	48	50.0	709	2	Q7KVBI_DROME
15	48	50.0	743	2	Q9WOC5_DROME
16	48	50.0	1081	1	KOKO_YEAST
17	48	50.0	1296	2	Q6CUE2_KLULA
18	47	49.0	248	2	Q5L2K2_STRT1
19	47	49.0	248	2	Q5M455_STRT2
20	47	49.0	254	2	Q6NJC6_CORDI
21	47	49.0	264	2	Q4NCZ5_9MICC
22	47	49.0	355	2	Q8R0F4_MOUSE
23	47	49.0	421	2	Q6N9D6_RHOPA
24	47	49.0	676	2	Q9C2A6_NEUCR
25	47	49.0	779	1	GSHAB_DESPS
26	47	49.0	1296	2	Q752D9_ASHGO
27	47	49.0	1976	2	Q4IP27_GIBZE
28	46	47.9	159	2	Q6KAP4_ORYSA
29	46	47.9	167	2	Q7SAV0_NEUCR
30	46	47.9	265	2	Q6CY53_KLULA
31	46	47.9	830	2	Q7QHW5_ANOGA

32	46	47.9	1222	1	WNK4_RAT
33	46	47.9	1292	2	Q6C910_YARLI
34	45	46.9	89	2	Q9SB08_ORYLO
35	45	46.9	90	2	Q9S7V3_ORYLO
36	45	46.9	90	2	Q9SB09_ORYLO
37	45	46.9	91	2	Q9ZWT5_ORYLO
38	45	46.9	150	2	Q5T558_HUMAN
39	45	46.9	197	2	Q5JM49_ORYSA
40	45	46.9	218	2	Q7M7M9_WOLSU
41	45	46.9	415	2	Q5WS17_LEGPA
42	45	46.9	449	2	Q89B44_BRAJA
43	45	46.9	488	1	KATL1_MOUSE
44	45	46.9	488	1	KATL1_RAT
45	45	46.9	490	1	KATL1_HUMAN
46	45	46.9	766	1	PKH1_YEAST
47	45	46.9	778	1	GSHAB_CLOPE
48	45	46.9	848	2	Q56VR4_MOUSE
49	45	46.9	881	2	Q4NCP6_ASPTU
50	45	46.9	948	2	Q6TAF9_SOLTU
51	45	46.9	965	2	Q6FTB4_CANGA
52	45	46.9	1074	2	Q6SGU6_BACLD
53	45	46.9	1222	1	WNK4_MOUSE
54	45	46.9	1222	2	Q4VAC1_MOUSE
55	45	46.9	1241	2	Q511H2_ENTHI
56	45	46.9	1288	2	Q7QAJ3_ANOGA
57	45	46.9	1338	2	Q9BKJ2_DROME
58	45	46.9	1674	2	Q9W0R8_DROME
59	44.5	46.4	362	2	Q4LGF7_9BURK
60	44.5	46.4	911	2	Q23629_CAEBL
61	44	45.8	115	2	Q8IPX9_DROME
62	44	45.8	125	2	Q4PCU2_USTWA
63	44	45.8	216	2	Q8XYP1_RALSO
64	44	45.8	251	2	Q4JTW6_CORJK
65	44	45.8	352	2	Q9P671_NEUCR
66	44	45.8	366	2	Q6SH28_9BACT
67	44	45.8	448	2	Q5KRG2_CRYNE
68	44	45.8	457	2	Q5SN33_CRYNE
69	44	45.8	552	2	Q86Z11_NEUCR
70	44	45.8	608	2	Q5LHS6_BACFN
71	44	45.8	608	2	Q64VS8_BACFR
72	44	45.8	609	2	Q8A532_BACTN
73	44	45.8	675	2	Q6BCL1_MOUSE
74	44	45.8	988	1	RG44_SOLBU
75	44	45.8	1044	2	Q9RK44_STRCO
76	44	45.8	1305	2	Q6FNZ8_CANGA
77	44	45.8	1312	1	RAD50_YEAST
78	44	45.8	1332	2	Q5AJE3_CANAL
79	44	45.8	2849	2	Q8W6K4_9CAUD
80	43.5	45.3	501	2	Q4J151_AZQVI
81	43.5	45.3	561	2	Q7UTG9_RHOSA
82	43	44.8	117	2	Q8XYS0_RALSO
83	43	44.8	148	2	Q4NLX7_9MICC
84	43	44.8	197	2	Q9ZY54_RHIME
85	43	44.8	258	2	Q7NWX5_CHRVO
86	43	44.8	273	2	Q4NTT3_9DELT
87	43	44.8	320	2	Q7W235_BORPA
88	43	44.8	320	2	Q7WR13_BORBR
89	43	44.8	335	2	Q8R002_BORBR
90	43	44.8	335	2	Q45375_BORPE
91	43	44.8	338	2	Q86K79_DICDI
92	43	44.8	343	2	Q550A8_DICDI
93	43	44.8	356	2	Q88NH5_PSEPK
94	43	44.8	356	2	Q82B48_STRAW
95	43	44.8	360	2	Q5QVC9_IDILO
96	43	44.8	373	2	Q4TWJ8_9SPHN
97	43	44.8	380	2	Q86F58_SCHJA
98	43	44.8	406	2	Q6H16_SCHJA
99	43	44.8	435	2	Q67J58_SYMTH
100	43	44.8	450	2	Q95NR5_9EUKA
101	43	44.8	450	2	Q95YH1_9EUKA
102	43	44.8	462	2	Q4MTJ6_THRPA
103	43	44.8	548	2	Q8R0S2_MOUSE
104	43	44.8	568	2	Q5N3W4_SYNPE

Q7tpk6	rattus norv
Q6c910	yarrowia li
Q9sb08	oryza longi
Q9s7v3	oryza longi
Q9sb09	oryza longi
Q9zwt5	oryza longi
Q5t558	homo sapien
Q5jm49	oryza sativ
Q7m7m9	wolinnella s
Q5ws17	legionella
Q89b44	bradyrhizob
Q8k0e4	mus musculus
Q5xik7	rattus norv
Q9bw62	homo sapien
Q03407	saccharomyc
Q8xk30	c glutathio
Q56vr4	mus musculus
Q4ncp6	aspergillus
Q6tbf9	solanum tub
Q6ftb4	candida gla
Q6sgu6	bacillus li
Q80ue6	mus musculus
Q4vac1	mus musculus
Q511h2	entamoeba h
Q7qaj3	anopheles g
Q9bkj2	drosophila
Q9w0r8	drosophila
Q4lcf7	burkholderi
Q23629	caenorhabdi
Q8ipx9	drosophila
Q4pcu2	ustilago ma
Q8xyp1	ralstonia s
Q4jtw6	corynebacte
Q9p671	neutrospora
Q6sh28	uncultured
Q5k92	cryptococcu
Q55n33	cryptococcu
Q86z11	neutrospora
Q5lha6	bacteroides
Q64vs8	bacteroides
Q8a532	bacteroides
Q6bcl1	mus musculus
Q7x339	solanum bul
Q9rk44	streptomyces
Q6fnz8	candida gla
P12753	saccharomyc
Q5ajc3	candida alb
Q8w6k4	sinorhizobi
Q4j151	azotobacter
Q7utg9	rhodospirell
Q8xys0	ralstonia s
Q4nlx7	arthrobacte
Q9zy54	rhizobium m
Q7nwx5	chromobacte
Q4ntt3	anaeromycob
Q7w235	bordetella
Q7wr13	bordetella
Q8r002	bordetella
Q45375	bordetella
Q86k79	dicytyosteli
Q550a8	dicytyosteli
Q88nh5	pseudomonas
Q82b48	streptomyces
Q5qvc9	idimarina
Q4tmj8	erythrobact
Q86f58	schistosoma
Q5dh16	schistosoma
Q67j58	symbiobacte
Q95nr5	pseudotrich
Q95yh1	pseudotrich
Q4mj36	thellieria p
Q8r0s2	mus musculus
Q5n3w4	synchococc

105	43	44.8	589	1	SYT_MYCPU	Q98ph6 mycoplasma	178	42	43.8	1748	1	ZOI_HUMAN	Q07157 homo sapien
106	43	44.8	674	2	Q4JUK6 CORJK	Q4juk6 corynebacte	179	42	43.8	1748	2	Q4ZGJ6 HUMAN	Q4zgj6 homo sapien
107	43	44.8	736	2	Q8T9Z2 DROME	Q8t9z2 drosophila	180	42	43.8	1906	2	Q9VJY7 DROME	Q9vjy7 drosophila
108	43	44.8	746	1	NAS30_CABEL	Q9n2v2 caenorhabdi	181	42	43.8	3680	2	Q5D1S7_PSEAE	Q5di17 pseudomonas
109	43	44.8	814	2	Q96D85 HUMAN	Q96d85 homo sapien	182	42	43.8	3680	2	Q5DIU0_PSEAE	Q5diu0 pseudomonas
110	43	44.8	843	2	Q94863 HUMAN	Q94863 homo sapien	183	42	43.8	4367	2	Q5DIS9_PSEAE	Q5dis9 pseudomonas
111	43	44.8	845	2	Q6BRD1_DEBHA	Q6brd1 debaryomyce	184	42	43.8	4372	2	Q5DIU1_PSEAE	Q5diu1 pseudomonas
112	43	44.8	963	2	Q6BRN0_HUMAN	Q6brn0 homo sapien	185	41.5	43.2	167	1	Y1358_VIBVY	Q7ml9 vibrio vuln
113	43	44.8	1076	2	Q6F0B9_CANGA	Q6f0b9 candida gla	186	41.5	43.2	183	2	Q8MWC3_PENMO	Q8mwc3 penaeus mon
114	43	44.8	1096	2	Q9DSU4_MOUSE	Q9dsu4 mus musculu	187	41.5	43.2	395	2	Q7UUE3_RHOBA	Q7uue3 rhodopirell
115	43	44.8	1114	2	P89460_HHV2	P89460 human herpe	188	41.5	43.2	503	1	CP3AC_CANFA	P24463 canis famill
116	43	44.8	1207	2	Q9DSV1_MOUSE	Q9dsv1 mus musculu	189	41.5	43.2	944	2	O5A3P6_CANAL	O5a3p6 candida alb
117	43	44.8	1243	1	WNK4_HUMAN	Q96j92 homo sapien	190	41.5	43.2	1153	2	O5A3V9_CANAL	O5a3v9 candida alb
118	42.5	44.3	515	2	Q74HQ4_LACJO	Q74hq4 lactobacill	191	41.5	43.2	1740	2	Q4SID5_TETNG	Q4sid5 tetraodon n
119	42.5	44.3	921	2	Q9VLR7_DROME	Q9vlr7 drosophila	192	41	42.7	29	2	Q8BTD7_MOUSE	Q8btd7 mus musculu
120	42	43.8	66	2	Q6IHG4_DROME	Q6ihg4 drosophila	193	41	42.7	78	2	Q9K787_BACHD	Q9k787 bacillus ha
121	42	43.8	84	2	Q9S7E0_ORYRU	Q9s7e0 oryza rufip	194	41	42.7	104	2	Q4LYP2_9BURK	Q4lyp2 burkholderi
122	42	43.8	84	2	Q9S7E1_ORYYZ	Q9s7e1 oryza merid	195	41	42.7	127	2	O52AS9_MAGGR	O52as9 magnaporthe
123	42	43.8	84	2	Q9SXL0_ORYGL	Q9sxl0 oryza glabe	196	41	42.7	133	2	Q7P090_CHRVO	Q7p090 chromobacte
124	42	43.8	84	2	Q9XJ50_ORYNI	Q9xj50 oryza niavar	197	41	42.7	139	2	Q7KX00_CABEL	Q7kx00 caenorhabdi
125	42	43.8	85	2	Q9XJ49_ORYYZ	Q9xj49 oryza barth	198	41	42.7	152	2	Q9ABN6_CAUCR	Q9abn6 caulobacter
126	42	43.8	86	2	Q9XJ48_ORYGL	Q9xj48 oryza glabe	199	41	42.7	166	2	Q4UFU0_XANCP	Q4upu0 xanthomonas
127	42	43.8	87	2	Q9XJ47_ORYYZ	Q9xj47 leersia per	200	41	42.7	166	2	Q8P490_XANCP	Q8p490 xanthomonas
128	42	43.8	89	2	Q9S7V2_ORYRU	Q9s7v2 oryza rufip	201	41	42.7	175	2	Q92LU1_RHIME	Q92lu1 rhizobium m
129	42	43.8	89	2	Q9ZWT6_ORYLO	Q9zwt6 oryza longi	202	41	42.7	177	2	Q945W8_ORYSA	Q945w8 oryza sativ
130	42	43.8	90	2	Q9SXL1_ORYSA	Q9sxl1 oryza sativ	203	41	42.7	177	2	Q8DKE2_SYNEL	Q8dke2 synecococc
131	42	43.8	90	2	Q9XJ46_ORYYZ	Q9xj46 oryza coarc	204	41	42.7	189	2	Q94K37_ARATH	Q94k37 arabidopsis
132	42	43.8	91	2	Q7GD96_ORYSA	Q7gd96 oryza sativ	205	41	42.7	195	2	Q8XVJ4_RALSO	Q8xvj4 ralstonia g
133	42	43.8	91	2	Q9S7S0_ORYSA	Q9s7s0 oryza sativ	206	41	42.7	214	2	Q6TR70_9STRA	Q6tr70 pythium aff
134	42	43.8	167	2	Q4TI94_TETNG	Q4ti94 tetraodon n	207	41	42.7	214	2	Q6TR71_9STRA	Q6tr71 pythium mid
135	42	43.8	208	2	O5YMF1_NOCPA	O5ymf1 nocardia fa	208	41	42.7	214	2	Q6TR72_9STRA	Q6tr72 pythium mon
136	42	43.8	239	2	O6EEM5_TRITU	O6eem5 triticum ae	209	41	42.7	223	1	VG32_BPMD2	O64226 mycobacteri
137	42	43.8	242	2	O6EEM3_TRITU	O6eem3 triticum tu	210	41	42.7	234	2	Q9SPD4_9STRA	Q9spd4 pythium aph
138	42	43.8	255	2	Q8S7S9_ORYSA	Q8s7s9 oryza sativ	211	41	42.7	235	2	Q8L4V6_ORYSA	Q8l4v6 oryza sativ
139	42	43.8	274	2	O6EEX0_WHEAT	O6eex0 triticum ae	212	41	42.7	270	2	Q4H6C5_9DEIO	Q4h6c5 deinococcus
140	42	43.8	275	2	O6EEM2_9POAL	O6eem2 triticum tu	213	41	42.7	279	1	PSA1_DROME	P12881 drosophila
141	42	43.8	275	2	O6EEM6_TRITU	O6eem6 triticum tu	214	41	42.7	288	2	O5UF67_9PROT	O5uff67 uncultured
142	42	43.8	275	2	O6EEX1_WHEAT	O6eex1 triticum ae	215	41	42.7	289	1	DAPA_CHLCV	O822h0 chlamydomo
143	42	43.8	275	1	GDBB_WHEAT	P06659 triticum ae	216	41	42.7	289	1	Q9W4H1_DROME	Q9w4h1 drosophila
144	42	43.8	311	2	Q9AEM7_CAUCR	Q9abm7 caulobacter	217	41	42.7	301	2	Q9K9Q8_BACHD	Q9k9q8 bacillus ha
145	42	43.8	316	2	Q4TBQ7_TETNG	Q4tbq7 tetraodon n	218	41	42.7	327	2	Q9U067_9ASIL	Q9u067 theriidae
146	42	43.8	365	2	Q9KXK4_STROO	Q9kxk4 streptomyce	219	41	42.7	331	2	Q89MP4_BRAJA	Q89mp4 bradyrhizob
147	42	43.8	366	2	Q4WOL2_ASPGIL	Q4wql2 aspergillus	220	41	42.7	335	2	Q7SLG5_NEUCR	Q7slg5 neucospora
148	42	43.8	369	2	Q4IRZ7_GIBZE	Q4ier7 gibberella	221	41	42.7	337	2	PM03_MOUSE	Q99n95 mus musculu
149	42	43.8	438	2	Q46124_HAECCO	Q46124 haemonchus	222	41	42.7	348	1	Q91XB2_MOUSE	Q91xb2 mus musculu
150	42	43.8	438	2	Q67BC0_9BILA	Q67bc0 cooperia on	223	41	42.7	348	2	Q74BV9_GEOSL	Q74bv9 geobacter s
151	42	43.8	491	1	CATA1_ORYSA	P29611 oryza sativ	224	41	42.7	358	2	Q4V8K5_RAT	Q4v8k5 rattus norv
152	42	43.8	492	2	Q6Z7B2_ORYSA	Q6z7b2 oryza sativ	225	41	42.7	411	2	O59713_PSEPU	O59713 pseudomonas
153	42	43.8	495	2	Q7UG67_RHOBA	Q7ug67 rhodopirell	226	41	42.7	421	2	Q4WW15_ASPFU	Q4ww15 aspergillus
154	42	43.8	504	2	Q9C148_COLGL	Q9c148 colletotric	227	41	42.7	446	2	O5K25_CRYNE	O5sk25 cryptococcu
155	42	43.8	548	2	Q5AUA2_EMENI	Q5aua2 aspergillus	228	41	42.7	448	2	Q9RDZ5_LEGPN	Q9rdz5 legionella
156	42	43.8	560	2	Q7UY15_RHOBA	Q7uy15 rhodopirell	229	41	42.7	474	2	O5WYC5_LEGPL	O5wyc5 legionella
157	42	43.8	564	2	Q6NBD1_RHOPA	Q6nbd1 rhodopseudo	230	41	42.7	474	2	O5X6X6_LEGPA	O5x6x6 legionella
158	42	43.8	592	2	Q73PD2_TREDE	Q73pd2 treponema d	231	41	42.7	474	2	O5ZXF8_LEGPH	O5zx18 legionella
159	42	43.8	615	2	O53HC5_HUMAN	O53hc5 homo sapien	232	41	42.7	474	2	O7UYV4_RHOBA	Q7uyv4 rhodopirell
160	42	43.8	615	2	Q8TAP0_HUMAN	Q8tap0 homo sapien	233	41	42.7	479	2	Q88X89_PSESM	Q88x89 pseudomonas
161	42	43.8	615	2	Q9NUX3_HUMAN	Q9nux3 homo sapien	234	41	42.7	493	2	Q67KB3_SYMTH	Q67kb3 symbiobacte
162	42	43.8	681	1	MAOC_ECOLI	P77455 escherichia	235	41	42.7	495	2	Q7UP64_RHOBA	Q7up64 rhodopirell
163	42	43.8	702	2	Q4LEZ1_9ENTR	Q4lez1 klebsiella	236	41	42.7	510	2	Q6H7D0_ORYSA	Q6h7d0 oryza sativ
164	42	43.8	730	2	Q4ZHV6_MANGE	Q4zhv6 manduca sex	237	41	42.7	513	2	O95PU8_CHITE	Q95pu8 chironomus
165	42	43.8	831	1	NAPA_PAPRN	O56350 paracoccus	238	41	42.7	522	2	Q6DHU3_YLDV	Q6dh10 yaba-like d
166	42	43.8	937	2	Q6Z3H8_ORYSA	Q6z3h8 oryza sativ	239	41	42.7	522	2	O68CU5_HUMAN	O68cu5 homo sapien
167	42	43.8	947	2	Q6TAF8_SOLTU	Q6taf8 solanum tub	240	41	42.7	540	2	FA20A_HUMAN	Q96mk3 homo sapien
168	42	43.8	970	1	RGA2_SOLBU	Q7xbg9 solanum bul	241	41	42.7	541	1	Q92X63_RHIME	Q92x63 rhizobium m
169	42	43.8	1016	2	O5E0L9_VIBF1	Q5e0l9 vibrio fisc	242	41	42.7	544	2	Q4UW05_XANCP	Q4uw05 xanthomonas
170	42	43.8	1025	1	RGAL_SOLBU	Q7xa42 solanum bul	243	41	42.7	560	2	Q8P838_XANCP	Q8p838 xanthomonas
171	42	43.8	1033	2	Q753H7_ASHGO	Q753h7 ashbya goss	244	41	42.7	568	2	Q7M2S2_PHOIL	Q7mza2 photorhabdu
172	42	43.8	1218	2	Q4VSW4_RAT	Q4vsw4 rattus norv	245	41	42.7	588	2	Q977Y3_CLOAB	Q977y3 clostridium
173	42	43.8	1309	2	Q4PFM8_USTWA	Q4pfm8 ustilago ma	246	41	42.7	606	2	PRIM_MYCBO	P63963 mycobacteri
174	42	43.8	1322	2	Q4XIH0_ASPFU	Q4xih0 aspergillus	247	41	42.7	639	1	PRIM_MYCTU	P63962 mycobacteri
175	42	43.8	1380	2	Q4P7H1_USTWA	Q4p7h1 ustilago ma	248	41	42.7	639	1	PRIM_MYCLE	Q6cc92 mycobacteri
176	42	43.8	1547	2	Q4T638_TETNG	Q4t638 tetraodon n	249	41	42.7	642	1	Q73Y28_MYCPA	Q73y28 mycobacteri
177	42	43.8	1692	2	Q6MZU1_HUMAN	Q6mzul homo sapien	250	41	42.7	650	2		

251	41	42.7	656	2	Q4FTY6_LEIMA	Q4fyc6 leishmania	324	40.5	42.2	496	2	Q8P9D2_XANCP	Q8p9d2 xanthomonas
252	41	42.7	659	2	Q8X0A3_NEUCR	Q8x0a3 neurospora	325	40.5	42.2	810	2	Q5BE84_EMENI	Q5be84 aspergillus
253	41	42.7	681	2	Q9KX55_STRCO	Q9kx55 streptomyces	326	40.5	42.2	891	2	Q60LW0_CAEBR	Q60lw0 caenorhabdi
254	41	42.7	692	2	Q4RH66_TETNG	Q4rh66 tetraodon n	327	40.5	42.2	947	2	Q4TPS2_9SPHN	Q4tps2 erythroba
255	41	42.7	694	1	ABCB7_MOUSE	Abcb7 mus musculu	328	40.5	42.2	1055	2	Q55XG3_CRYNE	Q55xg3 cryptococcu
256	41	42.7	713	2	Q6MGF7_NEUCR	Q6mgf7 neurospora	329	40.5	42.2	1055	2	Q5XWH0_CRYNE	Q5xwh0 cryptococcu
257	41	42.7	721	1	YIJ1_YEAST	Yij1 yeast	330	40	41.7	83	2	Q9PLD5_HUMAN	Q9pld5 homo sapien
258	41	42.7	739	2	Q5C117_CRYHO	Q5c117 cryptospori	331	40	41.7	83	2	Q8K6G5_STRP3	Q8k6g5 streptococc
259	41	42.7	740	2	Q5CPH0_CRYPV	Q5cph0 cryptospori	332	40	41.7	96	1	GATC_LBPIC	Gatc leptospira
260	41	42.7	756	2	Q69900_STRCO	Q69900 streptomyce	333	40	41.7	96	1	GATC_LBPIC	Gatc leptospira
261	41	42.7	809	2	Q7SHC5_NEUCR	Q7shc5 neurospora	334	40	41.7	110	2	Q4J3N9_AZOVI	Q4j3n9 azotobacter
262	41	42.7	815	2	Q80SB9_BRARE	Q80sb9 brachydanio	335	40	41.7	118	2	Q7MWK5_PORGI	Q7mwk5 porphyromon
263	41	42.7	815	2	Q8A1P3_BRARE	Q8a1p3 brachydanio	336	40	41.7	121	2	Q5YZP0_NOCFA	Q5yzp0 nocardia fa
264	41	42.7	828	2	Q7AC56_ECO57	Q7ac56 escherichia	337	40	41.7	141	2	Q82F63_STRAW	Q82f63 streptomyce
265	41	42.7	838	2	Q8CVW4_ECOL6	Q8cvw4 escherichia	338	40	41.7	147	2	Q586G7_TRYPANOMA	Q586g7 trypanosoma
266	41	42.7	828	2	Q83QV0_SHIFL	Q83qv0 shigella fl	339	40	41.7	150	2	Q8PPB2_XANAC	Q8ppb2 xanthomonas
267	41	42.7	828	2	Q8XB47_ECO57	Q8xb47 escherichia	340	40	41.7	156	2	Q6SZN5_RHOPA	Q6szn5 rhodococcus
268	41	42.7	831	2	Q88111_RHOSH	Q88111 rhodobacter	341	40	41.7	157	2	Q7LJP8_CAEBL	Q7ljp8 caenorhabdi
269	41	42.7	844	2	Q5EBA5_RAT	Q5eba5 rattus norv	342	40	41.7	161	2	Q8C3Z5_MOUSE	Q8c3z5 mus musculu
270	41	42.7	852	2	Q5B9M2_EMENI	Q5b9m2 aspergillus	343	40	41.7	176	2	Q6OCN6_METCA	Q6ocn6 methylococc
271	41	42.7	863	2	Q6D5N8_ERWCT	Q6d5n8 erwinia car	344	40	41.7	210	2	Q9KY49_STRCO	Q9ky49 streptomyce
272	41	42.7	880	2	Q4NU98_9DELT	Q4nu98 anaeromyxob	345	40	41.7	220	2	Q5T4W7_HUMAN	Q5t4w7 homo sapien
273	41	42.7	901	2	Q9HCE2_HUMAN	Q9hce2 homo sapien	346	40	41.7	220	2	Q89JU3_BRAJA	Q89ju3 bradyrhizob
274	41	42.7	909	2	Q61AS5_CAEBR	Q61as5 caenorhabdi	347	40	41.7	225	2	Q4T083_TETNG	Q4t083 tetraodon n
275	41	42.7	919	2	Q9BUM8_HUMAN	Q9bum8 homo sapien	348	40	41.7	231	2	Q63UD5_BURPS	Q63ud5 burkholderi
276	41	42.7	919	2	Q5R7H0_PONPY	Q5r7h0 pongo pygma	349	40	41.7	235	2	Q83S19_SHIFL	Q83s19 shigella fl
277	41	42.7	919	2	Q5R7H0_PONPY	Q5r7h0 pongo pygma	350	40	41.7	237	2	Q4FO73_9GAMM	Q4fo73 psychrobact
278	41	42.7	947	1	RG33_SOLBU	Q7xa40 solanum bul	351	40	41.7	240	1	NFSA_ECOLI	E17117 escherichia
279	41	42.7	954	2	Q75UQ6_ACHTE	Q75uq6 achaearenea	352	40	41.7	240	1	NFSA_SALTY	Q9z522 salmonella
280	41	42.7	965	2	Q54CR1_DICDI	Q54cr1 dictyosteli	353	40	41.7	240	2	Q7UD78_SHIFL	Q7ud78 shigella fl
281	41	42.7	985	2	Q52BA8_MAGGR	Q52ba8 magnaporthe	354	40	41.7	240	2	Q5FGP2_SALPA	Q5fgp2 salmonella
282	41	42.7	990	2	Q5KA12_CRYNE	Q5ka12 cryptococcu	355	40	41.7	240	2	Q8FJF1_ECOL6	Q8fjf1 escherichia
283	41	42.7	996	2	Q59LN7_CANAL	Q59ln7 candida alb	356	40	41.7	240	2	Q8Z848_SALTI	Q8z848 salmonella
284	41	42.7	1028	1	CARB_THET2	Q96495 thermus the	357	40	41.7	240	2	Q8X6S1_ECO57	Q8x6s1 escherichia
285	41	42.7	1028	2	Q5SKN1_THET8	Q5skn1 thermus the	358	40	41.7	249	2	Q9YD68_AERPE	Q9yad8 aeropyrum p
286	41	42.7	1032	2	Q9JY93_MOUSE	Q9jy93 mus musculu	359	40	41.7	254	2	Q4H858_9DEIO	Q4h858 deinococcus
287	41	42.7	1043	2	Q6NXZ1_MOUSE	Q6nxz1 mus musculu	360	40	41.7	263	2	Q4SKA1_TETNG	Q4ska1 tetraodon n
288	41	42.7	1059	2	Q5KG17_CRYNE	Q5kg17 cryptococcu	361	40	41.7	264	2	Q9S7E3_SOYBN	Q9s7e3 glycine max
289	41	42.7	1059	2	Q55S85_CRYNE	Q55s85 cryptococcu	362	40	41.7	274	2	Q4HVA2_GIBZE	Q4hva2 gibberella
290	41	42.7	1070	2	Q4SXN3_TETNG	Q4sxn3 tetraodon n	363	40	41.7	274	2	Q6ACQ2_LEIXX	Q6acq2 leifeonia x
291	41	42.7	1084	2	Q524P9_MAGGR	Q524p9 magnaporthe	364	40	41.7	312	2	P72585_SYNY3	P72585 synechocyst
292	41	42.7	1098	2	Q69Z85_MOUSE	Q69z85 mus musculu	365	40	41.7	321	2	Q4XTN7_PLACH	Q4xtn7 plasmodium
293	41	42.7	1116	2	Q4NXZ5_9DELT	Q4nxz5 anaeromyxob	366	40	41.7	323	2	Q8G6D9_BIFLO	Q8g6d9 bifidobacte
294	41	42.7	1142	2	Q4P848_USTMA	Q4p848 ustilago ma	367	40	41.7	324	2	Q8UAW4_AGR75	Q8uaw4 agrobacteri
295	41	42.7	1185	2	Q6K3C7_ORYSA	Q6k3c7 oryza sativ	368	40	41.7	328	2	Q97EW3_CLOAB	Q97ew3 clostridium
296	41	42.7	1214	2	Q9M1P1_ARATH	Q9m1p1 arabidopsis	369	40	41.7	341	2	Q5VOT4_HALMA	Q5vot4 haloarcula
297	41	42.7	1241	2	Q85444_9REOV	Q85444 diadromus p	370	40	41.7	347	2	Q9BRZ6_HUMAN	Q9brz6 homo sapien
298	41	42.7	1247	2	Q5WPB2_BRARE	Q5wpb2 brachydanio	371	40	41.7	353	2	Q80449_ARATH	Q80449 arabidopsis
299	41	42.7	1290	1	RAD50_SCHPO	Q9utj8 schizosacch	372	40	41.7	353	2	Q8LCL0_ARATH	Q8lcl0 arabidopsis
300	41	42.7	1294	2	Q7SHS5_NEUCR	Q7shs5 neurospora	373	40	41.7	356	2	Q4W78_PSSSY	Q4w78 pseudomonas
301	41	42.7	1318	2	Q4WUR8_ASPFU	Q4wur8 aspergillus	374	40	41.7	356	2	Q8YV53_PSSSM	Q8yv53 pseudomonas
302	41	42.7	1372	2	Q4SBL9_TETNG	Q4sbl9 tetraodon n	375	40	41.7	357	2	Q93ZC0_ARATH	Q93zc0 arabidopsis
303	41	42.7	1400	2	Q5KAY6_CRYNE	Q5kay6 cryptococcu	376	40	41.7	360	2	Q5F6X4_AZOSE	Q5f6x4 azoarcus sp
304	41	42.7	1424	2	Q5RGA8_BRARE	Q5rga8 brachydanio	377	40	41.7	366	2	Q5WA95_ORYSA	Q5wa95 oryza sativ
305	41	42.7	4547	2	Q52US9_OMPOL	Q52us9 omphalotus	378	40	41.7	366	2	Q9SXX9_ARATH	Q9sxx9 arabidopsis
306	41	42.7	7192	2	Q5Z1N5_NOCFA	Q5z1n5 nocardia fa	379	40	41.7	390	2	Q5Z807_ORYSA	Q5z807 oryza sativ
307	40.5	42.2	238	2	Q6ZAJ5_ORYSA	Q6zaj5 oryza sativ	380	40	41.7	397	2	Q8BWT3_MOUSE	Q8bwt3 mus musculu
308	40.5	42.2	260	1	NPD_HALMA	Q5v4g5 haloarcula	381	40	41.7	411	2	Q8K2Q7_MOUSE	Q8k2q7 mus musculu
309	40.5	42.2	300	2	Q5GXR9_XANOR	Q5gx9 xanthomonas	382	40	41.7	411	2	Q8BWM8_MOUSE	Q8bwm8 mus musculu
310	40.5	42.2	357	2	Q7YVH5_MONBE	Q7yv55 monosiga br	383	40	41.7	413	2	Q4J412_AZOVI	Q4j412 azotobacter
311	40.5	42.2	379	2	Q7V494_PROMM	Q7v494 prochlorococ	384	40	41.7	422	2	Q8K0F9_MOUSE	Q8k0f9 mus musculu
312	40.5	42.2	472	2	Q5ZC05_ORYSA	Q5zc05 oryza sativ	385	40	41.7	441	2	Q7S9E3_NEUCR	Q7s9e3 neurospora
313	40.5	42.2	478	2	Q5QUH7_IDILO	Q5quh7 idiomarina	386	40	41.7	446	2	Q5WPT2_LUTLO	Q5wpt2 lutomyia l
314	40.5	42.2	485	2	Q6LTP6_PHOPR	Q6ltp6 photobacter	387	40	41.7	462	2	Q86Y97_HUMAN	Q86y97 homo sapien
315	40.5	42.2	488	2	Q30862_VIBCH	Q30862 vibrio chol	388	40	41.7	464	2	Q92T12_RHIME	Q92t12 rhizobium m
316	40.5	42.2	488	2	Q9KQ66_VIBCH	Q9kq66 vibrio chol	389	40	41.7	474	2	Q9FAV0_9FLAO	Q9fav0 salegentiba
317	40.5	42.2	493	2	Q5E3P5_VIBF1	Q5e3p5 vibrio fisc	390	40	41.7	484	2	Q9FAX5_CAPGI	Q9fax5 capnocytoph
318	40.5	42.2	493	2	Q7X484_VIBF1	Q7x484 vibrio fisc	391	40	41.7	484	2	Q72W8_DESVH	Q72w8 desulfovibr
319	40.5	42.2	494	2	Q5GZM8_XANOR	Q5gzm8 xanthomonas	392	40	41.7	492	1	SYK_THETH	P41355 thermus the
320	40.5	42.2	494	2	Q8PLJ39_XANAC	Q8plj39 xanthomonas	393	40	41.7	492	2	Q5SUG7_THET8	Q5sug7 thermus the
321	40.5	42.2	496	1	C330_FUNHE	Q9pve8 fundulus he	394	40	41.7	492	2	Q72J19_THET2	Q72j19 thermus the
322	40.5	42.2	496	1	C356_FUNHE	Q8axy5 fundulus he	395	40	41.7	497	2	Q652F9_ORYSA	Q652f9 oryza sativ
323	40.5	42.2	496	2	Q4UUG5_XANCP	Q4uug5 xanthomonas	396	40	41.7	500	2	Q857X9_9CAUD	Q857x9 mycobacteri

397	40	41.7	508	2	Q6Z2J3_ORYSA	Q6Z2J3_oryza sativ	470	39.5	41.1	466	2	Q5KL83_CRYNE	Q5KL83_cryptococcu
398	40	41.7	521	1	PGS1_SACPS	P79001_saccharomyc	471	39.5	41.1	498	2	Q6BUV5_DEBHA	Q6BUV5_debaryomyce
399	40	41.7	536	2	Q8KJN9_RHILO	Q8KJN9_rhizobium l	472	39.5	41.1	502	2	Q4T616_TETNG	Q4T616_tetradodon n
400	40	41.7	536	2	Q89J12_RHILO	Q89J12_rhizobium l	473	39.5	41.1	543	2	O18556_MUSDO	O18556_musca domes
401	40	41.7	538	2	Q8PJ15_XANAC	Q8PJ15_xanthomonas	474	39.5	41.1	811	2	Q52339_NOCPA	Q52339_nocardia fa
402	40	41.7	544	2	Q4RXQ9_TETNG	Q4RXQ9_tetradodon n	475	39.5	41.1	811	2	Q51QE2_MAGGR	Q51QE2_magnaporthe
403	40	41.7	557	2	Q9FLK3_ARATH	Q9FLK3_arabidopsis	476	39.5	41.1	883	2	Q9LBD7_ARATH	Q9LBD7_arabidopsis
404	40	41.7	579	2	Q5Z2U9_MYCHV	Q5Z2U9_mycoplasma	477	39.5	41.1	1004	2	Q755Z3_ASHGO	Q755Z3_ashbya goss
405	40	41.7	612	2	Q6KAW3_PSEST	Q6KAW3_pseudomonas	478	39	40.6	29	2	Q9S8A7_SECCO	Q9S8A7_secale cere
406	40	41.7	617	1	ILVD_STRAW	Q8EWA3_streptomyc	479	39	40.6	78	2	Q4W4C1_ONCMY	Q4W4C1_oncorhynch
407	40	41.7	617	1	ILVD_STRCO	Q8E198_streptomyc	480	39	40.6	98	2	Q63M43_BURPS	Q63M43_burkholderi
408	40	41.7	620	2	Q7NRK83_GLOVI	Q7NRK83_gloeobacter	481	39	40.6	104	2	Q71VR6_LISMP	Q71VR6_listeria mo
409	40	41.7	634	2	Q7RG66_PLAYO	Q7RG66_plasmodium	482	39	40.6	108	1	KV1R_HUMAN	P01610_homo sapien
410	40	41.7	644	2	Q5AVL4_BRENT	Q5AVL4_aspergillus	483	39	40.6	119	2	Q8TRX1_METAC	Q8TRX1_methanosarc
411	40	41.7	644	2	Q89E32_BRAJA	Q89E32_bradyrhizob	484	39	40.6	131	2	Q8BVE7_MOUSE	Q8BVE7_mus musculu
412	40	41.7	644	2	Q4SUJ8_TETNG	Q4SUJ8_tetradodon n	485	39	40.6	135	2	Q7NH51_GLOVI	Q7NH51_gloeobacter
413	40	41.7	651	2	Q4W4T2_TRYRA	Q4W4T2_trypanosoma	486	39	40.6	136	2	Q63VE6_BURPS	Q63VE6_burkholderi
414	40	41.7	665	1	LIN233_CAEL	Q09990_caenorhabdi	487	39	40.6	136	2	Q62FF8_BURMA	Q62FF8_burkholderi
415	40	41.7	680	2	Q96TX5_NEUCR	Q96TX5_neutrospora	488	39	40.6	161	2	Q8N910_HUMAN	Q8N910_homo sapien
416	40	41.7	687	2	Q7S7S9_NEUCR	Q7S7S9_neutrospora	489	39	40.6	161	2	Q8VEA3_MOUSE	Q8VEA3_mus musculu
417	40	41.7	695	2	Q70KH7_9ACTO	Q70KH7_streptomyc	490	39	40.6	166	2	Q5DU16_MOUSE	Q5DU16_mus musculu
418	40	41.7	698	2	Q8PXU6_METMA	Q8PXU6_methanosarc	491	39	40.6	179	2	Q6E850_CAMJE	Q6E850_campylobact
419	40	41.7	699	2	Q4IRI7_GIBZE	Q4IRI7_gibberella	492	39	40.6	188	2	Q6A842_PROAC	Q6A842_propionibac
420	40	41.7	702	2	Q5BD26_EMENI	Q5BD26_aspergillus	493	39	40.6	199	1	YIHX_SHIFL	P0A8Y3_escherichia
421	40	41.7	703	2	Q7SLP9_CAEL	Q7SLP9_caenorhabdi	494	39	40.6	199	1	YIHX_SHIFL	P0A8Y4_shigella fl
422	40	41.7	703	2	Q7LJP9_CAEL	Q7LJP9_plasmodium	495	39	40.6	202	2	Q7VVK7_BORPE	Q7VVK7_bordetella
423	40	41.7	705	2	Q8IKJ7_PLAF7	Q8IKJ7_plasmodium	496	39	40.6	202	2	Q7VVK7_BORPE	Q7VVK7_bordetella
424	40	41.7	715	2	Q61XG8_CAEBR	Q61XG8_caenorhabdi	497	39	40.6	203	2	Q5AX23_EMENI	Q5AX23_aspergillus
425	40	41.7	752	1	ABCB7_HUMAN	Q70488_rattus norv	498	39	40.6	205	2	Q6G3H8_BAHE	Q6G3H8_caenorhabdi
426	40	41.7	752	1	ABCB7_RAT	Q70488_rattus norv	499	39	40.6	206	2	Q622B1_CAEBR	Q622B1_caenorhabdi
427	40	41.7	774	2	Q4T491_TETNG	Q4T491_tetradodon n	500	39	40.6	206	2	Q8FBF4_ECOL6	Q8FBF4_escherichia
428	40	41.7	776	2	Q8LPH7_ARATH	Q8LPH7_arabidopsis	501	39	40.6	206	2	Q8X8D1_ECOL5	Q8X8D1_escherichia
429	40	41.7	784	2	Q5U400_BRARE	Q5U400_brachydanio	502	39	40.6	207	2	Q4UR26_XANCP	Q4UR26_xanthomonas
430	40	41.7	793	2	Q5B305_EMENI	Q5B305_aspergillus	503	39	40.6	207	2	Q63N03_BURPS	Q63N03_burkholderi
431	40	41.7	815	2	Q9SXY0_ARATH	Q9SXY0_arabidopsis	504	39	40.6	207	2	Q62AD9_BURMA	Q62AD9_burkholderi
432	40	41.7	817	2	Q8JG38_BRARE	Q8JG38_brachydanio	505	39	40.6	207	2	Q8PCF6_XANCP	Q8PCF6_xanthomonas
433	40	41.7	831	1	NAPA_RHOSH	Q53176_bordetacter	506	39	40.6	211	2	Q6LNY6_PHOPR	Q6LNY6_photobacter
434	40	41.7	835	2	Q5BA14_EMENI	Q5BA14_aspergillus	507	39	40.6	213	2	Q6K7Q2_ORYSA	Q6K7Q2_oryza sativ
435	40	41.7	838	2	Q6LQJ3_PHOPR	Q6LQJ3_photobacter	508	39	40.6	217	2	Q6ZTJ9_HUMAN	Q6ZTJ9_homo sapien
436	40	41.7	855	2	Q5F342_CHICK	Q5F342_gallus gall	509	39	40.6	218	2	Q60A68_METCA	Q60A68_methylococc
437	40	41.7	858	2	Q9FHF6_ARATH	Q9FHF6_arabidopsis	510	39	40.6	221	2	Q4HDA1_CAMCO	Q4HDA1_campylobact
438	40	41.7	885	2	Q4HGM1_9DEIO	Q4HGM1_deinococcus	511	39	40.6	225	2	Q9KEP4_BACHD	Q9KEP4_bacillus ha
439	40	41.7	914	2	Q5ATB5_EMENI	Q5ATB5_aspergillus	512	39	40.6	226	2	Q8BJN4_MOUSE	Q8BJN4_mus musculu
440	40	41.7	920	2	Q4KLP0_RAT	Q4KLP0_rattus norv	513	39	40.6	227	2	Q8YGB3_BRUME	Q8YGB3_brucella me
441	40	41.7	927	2	Q6P286_XENLA	Q6P286_xenopus lae	514	39	40.6	233	2	Q4LIQ6_9BURK	Q4LIQ6_burkholderi
442	40	41.7	948	2	Q6TAF7_9SOLN	Q6TAF7_solanum tar	515	39	40.6	233	2	Q8C2Z7_MOUSE	Q8C2Z7_mus musculu
443	40	41.7	1037	2	Q4T8U2_TETNG	Q4T8U2_tetradodon n	516	39	40.6	239	2	Q7WA04_BORFA	Q7WA04_bordetella
444	40	41.7	1145	2	Q9FLJ3_ARATH	Q9FLJ3_arabidopsis	517	39	40.6	243	2	Q9FTB9_AEGTA	Q9FTB9_aegilops ta
445	40	41.7	1146	2	Q8BXJ3_MOUSE	Q8BXJ3_mus musculu	518	39	40.6	243	2	Q9FTC0_AEGTA	Q9FTC0_aegilops ta
446	40	41.7	1160	2	Q6QX58_ARATH	Q6QX58_arabidopsis	519	39	40.6	243	2	Q8GC44_FLAS2	Q8GC44_flavobacter
447	40	41.7	1198	2	Q9C858_ARATH	Q9C858_arabidopsis	520	39	40.6	244	2	Q8L938_ARATH	Q8L938_arabidopsis
448	40	41.7	1202	2	Q9C608_ARATH	Q9C608_arabidopsis	521	39	40.6	244	2	Q9SSK4_ARATH	Q9SSK4_arabidopsis
449	40	41.7	1205	2	Q9FLJ3_ARATH	Q9FLJ3_arabidopsis	522	39	40.6	246	2	Q8W6F5_9CAUD	Q8W6F5_sinorhizobi
450	40	41.7	1205	2	Q7TNH0_MOUSE	Q7TNH0_mus musculu	523	39	40.6	250	2	Q9FS76_WHEAT	Q9FS76_triticum ae
451	40	41.7	1298	1	RAD50_CAEL	Q44139_caenorhabdi	524	39	40.6	251	1	GDB0_WHEAT	P08079_triticum ae
452	40	41.7	1298	2	Q9FHF3_ARATH	Q9FHF3_arabidopsis	525	39	40.6	252	2	Q9FS77_WHEAT	Q9FS77_triticum ae
453	40	41.7	1305	1	RPOC_UREPA	Q9PQV5_ureaplasma	526	39	40.6	254	2	Q9FS56_9POAL	Q9FS56_triticum va
454	40	41.7	1354	2	Q6OLV0_CAEBR	Q6OLV0_caenorhabdi	527	39	40.6	254	2	Q6E8W8_WHEAT	Q6E8W8_triticum ae
455	40	41.7	1462	2	Q5BG44_EMENI	Q5BG44_aspergillus	528	39	40.6	257	2	Q9FS62_WHEAT	Q9FS62_triticum ae
456	40	41.7	1483	2	Q6YNS3_DROME	Q6YNS3_drosophila	529	39	40.6	259	2	Q9FS75_WHEAT	Q9FS75_triticum ae
457	40	41.7	1483	2	Q8INT5_DROME	Q8INT5_drosophila	530	39	40.6	264	2	Q9FS74_WHEAT	Q9FS74_triticum ae
458	40	41.7	1802	2	Q7QCF0_ANOGA	Q7QCF0_anopheles g	531	39	40.6	267	2	Q9FEG9_WHEAT	Q9FEG9_triticum ae
459	40	41.7	2125	2	Q96057_HALRO	Q96057_holocynthia	532	39	40.6	267	2	Q9FS60_WHEAT	Q9FS60_triticum ae
460	40	41.7	2475	2	Q4IRX1_GIBZE	Q4IRX1_gibberella	533	39	40.6	270	2	Q9FS69_WHEAT	Q9FS69_triticum ae
461	40	41.7	2848	2	Q4N4Y7_THEPA	Q4N4Y7_theileria p	534	39	40.6	271	1	VIUB_VIBVU	Q56743_vibrio vuln
462	39.5	41.1	157	2	Q5FN11_GLUOX	Q5FN11_gloeobact	535	39	40.6	271	2	Q7MCT3_VIBVY	Q7MCT3_vibrio vuln
463	39.5	41.1	228	2	Q5F5W4_NEIG1	Q5F5W4_neisseria g	536	39	40.6	279	2	Q8L6B2_WHEAT	Q8L6B2_triticum ae
464	39.5	41.1	230	2	Q4XQN1_PLACH	Q4XQN1_plasmodium	537	39	40.6	279	2	Q8L6B3_WHEAT	Q8L6B3_triticum ae
465	39.5	41.1	244	2	Q945W3_ORYSA	Q945W3_oryza sativ	538	39	40.6	279	2	Q8L6B4_WHEAT	Q8L6B4_triticum ae
466	39.5	41.1	312	2	Q4UWC7_XANCP	Q4UWC7_xanthomonas	539	39	40.6	283	2	Q9FS73_WHEAT	Q9FS73_triticum ae
467	39.5	41.1	312	2	Q8P763_XANCP	Q8P763_xanthomonas	540	39	40.6	286	2	Q742F7_MYCPA	Q742F7_mycobacteri
468	39.5	41.1	463	2	Q599M9_9BETA	Q599M9_murid herpe	541	39	40.6	292	2	Q67PS1_SYMPH	Q67PS1_symbiobacte
469	39.5	41.1	466	2	Q55U29_CRYNE	Q55U29_cryptococcu	542	39	40.6	292	2	Q67PS1_SYMPH	Q67PS1_symbiobacte

543	39	40.6	293	2	Q8J326_BRARE	Q8J326 brachydanio	616	39	40.6	553	2	Q6KCW3_ECOLI	Q6kcw3 escherichia
544	39	40.6	295	2	Q9XEW0_WHEAT	Q9xew0 triticum ae	617	39	40.6	554	2	Q6WI92_BPKV4	Q6wi92 bacterioph
545	39	40.6	296	2	Q6B199_ECOLI	Q6b199 escherichia	618	39	40.6	555	2	Q5GQB5_9CAUD	Q5gqb5 bacterioph
546	39	40.6	298	2	Q9SYX8_WHEAT	Q9syx8 triticum ae	619	39	40.6	560	2	Q91O27_ADE02	Q91o27 human adeno
547	39	40.6	302	2	Q9ZIS4_ECOLI	Q9zis4 escherichia	620	39	40.6	571	2	Q9S2D8_ARATH	Q9s2d8 arabidopsis
548	39	40.6	305	2	Q9ZIT1_ECOLI	Q9zit1 escherichia	621	39	40.6	575	2	Q9VP62_DR0ME	Q9vp62 drosophila
549	39	40.6	306	2	Q71VW7_MUSAC	Q71vm7 musa acumin	622	39	40.6	596	2	Q9V5L3_BRAJA	Q9v5l3 bradyrhizob
550	39	40.6	308	2	Q8KMW3_ECOLI	Q8kmw3 escherichia	623	39	40.6	597	2	Q9W039_HUMAN	Q9w039 homo sapien
551	39	40.6	310	2	Q9ZIT9_ECOLI	Q9zit9 escherichia	624	39	40.6	611	2	Q6ACA4_LEIXX	Q6aca4 leifsonia x
552	39	40.6	311	2	Q4H5H8_9DEIO	Q4h5h8 deinococcus	625	39	40.6	611	2	Q7UV55_RHOBA	Q7uv55 rhodospirell
553	39	40.6	316	2	Q4H5H8_9DEIO	Q4h5h8 deinococcus	626	39	40.6	611	2	Q9VRF6_DR0ME	Q9vrf6 drosophila
554	39	40.6	319	1	RFAC_ECOLI	Q5yvt9 nocardia fa	627	39	40.6	611	2	Q7NDM4_GLOVI	Q7ndm4 gloeobacter
555	39	40.6	320	1	Q83J18_SHIFL	Q83j18 escherichia	628	39	40.6	635	1	MPP4_MOUSE	Q6p7f1 mus musculu
556	39	40.6	321	2	Q6S4E7_ORYSA	Q6s4e7 oryza sativ	629	39	40.6	637	2	Q53TT3_HUMAN	Q53tt3 homo sapien
557	39	40.6	326	2	Q8FC98_ECOLI	Q8fc98 escherichia	630	39	40.6	637	2	Q5BGY6_EMENI	Q5bgy6 aspergillus
558	39	40.6	330	2	Q846W0_STRCM	Q846w0 streptomyce	631	39	40.6	642	2	Q8BAP3_SHEON	Q8bap3 shewanella
559	39	40.6	330	2	Q8XDD3_ECO57	Q8xdd3 escherichia	632	39	40.6	645	2	Q8MS43_DR0ME	Q8ms43 drosophila
560	39	40.6	330	2	Q9QZL7_MOUSE	Q9qzl7 mus musculu	633	39	40.6	647	2	YDCP_ECOLI	Q8me43 drosophila
561	39	40.6	331	2	Q8L672_ORYSA	Q8l672 oryza sativ	634	39	40.6	653	1	Q761O4_ESCHRIC	Q761o4 escherichia
562	39	40.6	339	2	Q6LLM4_CAEBR	Q6llm4 caenorhabdi	635	39	40.6	659	2	Q76175_HUMAN	Q76175 homo sapien
563	39	40.6	340	2	Q5PMW1_SALPA	Q5pmw1 salmonella	636	39	40.6	659	2	Q8TE77_HUMAN	Q8te77 homo sapien
564	39	40.6	341	2	Q4HD93_CAMCO	Q4hd93 campylobact	637	39	40.6	659	2	Q6PK42_HUMAN	Q6pk42 homo sapien
565	39	40.6	341	2	Q89513_9VIRU	Q89513 andean pota	638	39	40.6	661	2	Q95JA4_PIG	Q95ja4 sus scrofa
566	39	40.6	351	2	Q5BDC2_EMENI	Q5bdc2 aspergillus	639	39	40.6	665	2	Q526T7_MAGGR	Q526t7 magnaporth
567	39	40.6	367	2	Q9VBX2_DR0ME	Q9vbx2 drosophila	640	39	40.6	665	2	Q9S814_HUMAN	Q9s814 homo sapien
568	39	40.6	368	2	Q6ESQ5_ORYSA	Q6esq5 oryza sativ	641	39	40.6	666	2	Q4QOL3_LEIMA	Q4qol3 leishmania
569	39	40.6	371	2	Q4LQ10_9BURK	Q4lq10 burkholderi	642	39	40.6	667	2	Q8ZUL8_PYRAE	Q8zul8 pyrobaculum
570	39	40.6	376	2	Q5CPH6_CRYPV	Q5cph6 cryptospori	643	39	40.6	667	2	Q7AE60_ECO57	Q7ae60 escherichia
571	39	40.6	376	2	Q5CFX4_CRYHO	Q5cfx4 cryptospori	644	39	40.6	667	2	Q8CW32_ECOLI	Q8cw32 escherichia
572	39	40.6	379	2	Q52KQ6_MOUSE	Q52kq6 mus musculu	645	39	40.6	667	2	Q8X9V9_ECO57	Q8x9v9 escherichia
573	39	40.6	380	1	HRB2_MOUSE	Q8bga5 mus musculu	646	39	40.6	670	1	ANGB1_HUMAN	Q8vcu0 mus musculu
574	39	40.6	380	2	Q5EBJ5_MOUSE	Q5ebj5 mus musculu	647	39	40.6	670	1	Q8NCS9_HUMAN	Q8ncs9 homo sapien
575	39	40.6	389	2	Q5DE82_SCHJA	Q5de82 schistosoma	648	39	40.6	674	2	Q8DL46_SYNEL	Q8dl46 synchococc
576	39	40.6	393	2	Q7XBY6_ORYSA	Q7xby6 oryza sativ	649	39	40.6	692	1	YY13_CABEL	Q18775 caenorhabdi
577	39	40.6	395	2	P70732_AZOBRA	Q70732 azospirillu	650	39	40.6	694	2	Q5CL30_SCHJA	Q5cl30 schistosoma
578	39	40.6	406	2	Q9RK30_STRCO	Q9rk30 streptomyce	651	39	40.6	731	1	GLGB_CORGL	Q8nr40 corynebacte
579	39	40.6	407	2	Q64OR3_MOUSE	Q64or3 mus musculu	652	39	40.6	755	2	Q82KQ8_STRAW	Q82kq8 streptomyce
580	39	40.6	413	2	Q4W15_PSEBSY	Q4w15 pseudomonas	653	39	40.6	770	2	Q413G5_GIBZE	Q413g5 gibberella
581	39	40.6	415	2	Q764M6_PIG	Q764m6 sus scrofa	654	39	40.6	772	2	Q59HH8_HUMAN	Q59hh8 homo sapien
582	39	40.6	419	2	Q6AGN2_LEIXX	Q6agn2 leifsonia x	655	39	40.6	776	2	Q7NP60_GLOVI	Q7np60 gloeobacter
583	39	40.6	424	2	ELK1_MOUSE	Q4p18 ustilago na	656	39	40.6	792	2	Q4Q175_LEIMA	Q4q175 leishmania
584	39	40.6	429	1	Q8LC35_ARATH	Q8lc35 arabidopsis	657	39	40.6	795	2	Q8J2W5_PYRAB	Q8j2w5 pyrococcus
585	39	40.6	429	2	Q9SAD0_ARATH	Q9sad0 arabidopsis	658	39	40.6	805	1	PK2L1_HUMAN	Q8j2w5 pyrococcus
586	39	40.6	429	2	Q9SAD0_ARATH	Q9sad0 arabidopsis	659	39	40.6	805	1	Q5W039_HUMAN	Q5w039 homo sapien
587	39	40.6	429	2	Q7TPR9_MOUSE	Q7tp9 mus musculu	660	39	40.6	827	2	Q91HX3_MACMG	Q91hx3 magnetospir
588	39	40.6	441	1	MPP4_RAT	Q9gyh1 rattus norv	661	39	40.6	837	2	Q8S9E5_BRAJA	Q8s9e5 bradyrhizob
589	39	40.6	452	2	Q9SGX8_ARATH	Q9sgx8 arabidopsis	662	39	40.6	838	2	Q9F5L3_BRAJA	Q9f5l3 bradyrhizob
590	39	40.6	467	1	IRF6_MOUSE	Q9p7431 mus musculu	663	39	40.6	841	2	Q6CDQ9_YARLI	Q6cdq9 yarrowia li
591	39	40.6	467	2	Q91VD0_MOUSE	Q91vd0 mus musculu	664	39	40.6	893	2	Q88EX2_PSEPK	Q88ex2 pseudomonas
592	39	40.6	468	2	Q8ZXM0_PYRAE	Q8zxm0 pyrobaculum	665	39	40.6	909	2	Q52EW8_MAGGR	Q52ew8 magnaporth
593	39	40.6	468	2	Q9SRF5_DR0ME	Q9srf5 drosophila	666	39	40.6	914	2	Q5R452_PONPY	Q5r452 pongo pygma
594	39	40.6	474	2	Q4P418_USTMA	Q4p418 ustilago na	667	39	40.6	915	1	SAFB1_HUMAN	Q15424 homo sapien
595	39	40.6	474	2	Q7UMB5_RHOBA	Q7umb5 rhodospirell	668	39	40.6	916	2	Q4R4L6_MACFA	Q4r4l6 macaca fasc
596	39	40.6	476	2	Q8MP11_CABEL	Q8mp11 caenorhabdi	669	39	40.6	931	1	SAFB1_RAT	Q88453 rattus norv
597	39	40.6	476	2	Q9BKT6_CABEL	Q9bkt6 caenorhabdi	670	39	40.6	941	2	Q8DFP9_VIBPU	Q8dfp9 vibrio vuln
598	39	40.6	480	2	Q5FRQ9_GLUOX	Q5frq9 gluconobact	671	39	40.6	941	2	Q87RF0_VIBPO	Q87rf0 vibrio para
599	39	40.6	484	2	Q8C157_MOUSE	Q8c157 mus musculu	672	39	40.6	941	2	Q7WMN5_VIBVY	Q7wmn5 vibrio vuln
600	39	40.6	488	2	Q9DGD7_ORYLA	Q9dgd7 oryzias lat	673	39	40.6	945	2	Q92Q30_RHIME	Q92q30 rhizobium m
601	39	40.6	488	2	Q5U3S1_BRARE	Q5u3s1 brachydanio	674	39	40.6	959	2	Q9VN15_DR0ME	Q9vn15 drosophila
602	39	40.6	499	2	Q5B713_EMENI	Q5b713 aspergillus	675	39	40.6	984	2	Q9XCD4_THFPU	Q9xcd4 thelmonos
603	39	40.6	513	2	Q8N9L8_HUMAN	Q8n9l8 homo sapien	676	39	40.6	988	2	Q4WPR1_ASFPU	Q4wpr1 aspergillus
604	39	40.6	521	1	C12D1_DR0ME	P82712 drosophila	677	39	40.6	992	2	Q5BBR3_EMENI	Q5bbr3 aspergillus
605	39	40.6	521	1	PGS1_VFAST	P25578 saccharomyc	678	39	40.6	992	2	Q9UV08_EMENI	Q9uv08 emericella
606	39	40.6	521	2	Q7XR10_DR0ME	Q7xr10 drosophila	679	39	40.6	997	2	Q6BK17_DEBHA	Q6bk17 debaryomyce
607	39	40.6	531	2	Q966P6_CABEL	Q966p6 caenorhabdi	680	39	40.6	1010	2	Q7RX84_NEUCR	Q7rx84 neurospora
608	39	40.6	531	2	Q52273_LACIA	Q52273 lactococcu	681	39	40.6	1014	2	Q6W5R1_9ACTO	Q6w5r1 streptomyce
609	39	40.6	531	2	Q54431_9LACT	Q54431 lactococcu	682	39	40.6	1017	2	Q6CLA7_KJULA	Q6cla7 klyveromyc
610	39	40.6	531	2	Q8KLM9_STRTR	Q8klm9 streptococc	683	39	40.6	1028	1	MYO1C_HUMAN	Q00159 homo sapien
611	39	40.6	531	2	Q9RNW1_STRTR	Q9rnw1 streptococc	684	39	40.6	1057	2	Q4WQ57_ASFPU	Q4wq57 aspergillus
612	39	40.6	537	2	Q7QC68_STRTR	Q7qc68 streptococc	685	39	40.6	1060	2	Q4QH32_LEIMA	Q4qh32 leishmania
613	39	40.6	537	2	Q7BLN9_LACIA	Q7bln9 lactococcu	686	39	40.6	1124	2	Q4QA08_LEIMA	Q4qa08 leishmania
614	39	40.6	537	2	Q9KIN3_9LACT	Q9kin3 lactococcu	687	39	40.6	1141	2	Q5B2L4_EMENI	Q5b2l4 aspergillus
615	39	40.6	553	1	MIS_RAT	P49000 rattus norv	688	39	40.6	1244	2	Q4P262_USTMA	Q4p262 ustilago ma

689	39	40.6	1254	1	UBP12 YEAST	P39538	saccharomyc	762	38	39.6	104	2	005263_BACSU	005263 bacillus su
690	39	40.6	1258	2	Q4G227_9TRYP	Q4G227 trypanosoma	763	38	39.6	106	2	Q94IU1_9ROSA	Q94IU1 fragaria nu	
691	39	40.6	1305	2	Q757A3_ASHGO	Q757A3 ashbya goss	764	38	39.6	106	2	Q94IU3_FRAVE	Q94IU3 fragaria ve	
692	39	40.6	1418	2	Q4WH7_ASPFU	Q4WH7 aspergillus	765	38	39.6	110	2	Q51TT1_MAGGR	Q51TT1 magnaporth	
693	39	40.6	1451	2	Q4RH5_TETNG	Q4RH5 tetraodon n	766	38	39.6	124	2	Q54II2_DICDI	Q54II2 dictyosteli	
694	39	40.6	1558	2	Q8II26_PLAF7	Q8II26 plasmodium	767	38	39.6	127	2	Q865F9_PIG	Q865F9 sus scrofa	
695	39	40.6	1631	2	Q6AUT8_ORYSA	Q6AUT8 oryza sativ	768	38	39.6	129	2	Q9XW8_CAEEL	Q9XW8 caenorhabdi	
696	39	40.6	1683	2	Q61FL3_CAEER	Q61FL3 caenorhabdi	769	38	39.6	131	1	PROF4_MAIZE	Q22655 zea mays (m	
697	39	40.6	1734	1	MAST2_MOUSE	Q60592 mus musculu	770	38	39.6	131	1	PROF5_MAIZE	Q9FR39 zea mays (m	
698	39	40.6	1798	1	MAST2_HUMAN	Q60048 homo sapien	771	38	39.6	135	2	Q767M9_PIG	Q767M9 sus scrofa	
699	39	40.6	1829	1	DPOL_THEST	Q33845 thermococcu	772	38	39.6	136	2	Q9YE11_AERPE	Q9YE11 aeropyrum p	
700	39	40.6	2179	1	K0310_HUMAN	Q15027 homo sapien	773	38	39.6	136	2	Q7RU15_NEUCR	Q7RU15 neurospora	
701	39	40.6	2195	2	Q4ST61_TETNG	Q45027 tetraodon n	774	38	39.6	142	2	Q6YRX1_SYNY3	Q6YRX1 synechocyst	
702	39	40.6	2333	2	Q615U0_CAEER	Q615U0 caenorhabdi	775	38	39.6	166	2	Q6NIV2_CORDI	Q6NIV2 corynebacte	
703	39	40.6	2361	2	Q4N0X8_THEPA	Q4N0X8 theileria p	776	38	39.6	168	2	Q52C06_MAGGR	Q52C06 magnaporth	
704	39	40.6	3171	2	Q7NUA0_CHRVO	Q7NUA0 chromobacte	777	38	39.6	170	2	Q4JTM2_CORJK	Q4JTM2 corynebacte	
705	39	40.6	3554	2	Q7NUA1_CHRVO	Q7NUA1 chromobacte	778	38	39.6	173	2	Q4PBT7_USTWA	Q4PBT7 ustilago ma	
706	39	40.6	4939	2	Q4QGE2_LEIMA	Q4QGE2 leishmania	779	38	39.6	177	1	PUR6_CORAM	Q44679 corynebacte	
707	39	40.6	6274	1	Q63UA4_BURPS	Q63UA4 burkholderi	780	38	39.6	179	2	Q8N4G4_HUMAN	Q8N4G4 homo sapien	
708	38.5	40.1	167	1	Y2912_VIBVU	Q8D8Q7 vibrio vuln	781	38	39.6	193	2	Q4H485_HUMAN	Q4H485 homo sapien	
709	38.5	40.1	220	2	Q5TMV6_ANOGA	Q5TMV6 anopheles g	782	38	39.6	196	2	Q6A733_PROAC	Q6A733 propionibac	
710	38.5	40.1	220	2	Q5TMV7_ANOGA	Q5TMV7 anopheles g	783	38	39.6	200	2	Q8G6X4_BIFLO	Q8G6X4 bifidobacte	
711	38.5	40.1	220	2	Q57CN5_BRUAB	Q57CN5 bruceella ab	784	38	39.6	206	1	GSTT_CAEEL	P31253 caenorhabdi	
712	38.5	40.1	220	2	Q8G057_BRUSU	Q8G057 bruceella su	785	38	39.6	206	2	Q8NDT9_HUMAN	Q8NDT9 homo sapien	
713	38.5	40.1	220	2	Q8YH08_BRUME	Q8YH08 bruceella me	786	38	39.6	206	2	Q8G4A7_BIFLO	Q8G4A7 bifidobacte	
714	38.5	40.1	259	2	Q5ZNZ7_VIRU	Q5ZNZ7 cotesia con	787	38	39.6	209	2	Q8XWB4_RALSO	Q8XWB4 ralsconia s	
715	38.5	40.1	329	1	IPNS_STRJU	P18286 streptomyce	788	38	39.6	210	2	Q62JL6_BURMA	Q62JL6 burkholderi	
716	38.5	40.1	339	1	NAMA_CLOAB	Q37866 clostridium	789	38	39.6	220	1	RK5_SPTOL	P82192 spinacia ol	
717	38.5	40.1	343	2	Q8G913_NITEU	Q8G913 nitrosomona	790	38	39.6	220	2	Q86030_HUMAN	Q86030 homo sapien	
718	38.5	40.1	410	2	Q8FU88_COREF	Q8FU88 corynebacte	791	38	39.6	220	2	Q81F84_DROME	Q81F84 drosophila	
719	38.5	40.1	420	2	Q6CKA4_KLJAB	Q6CKA4 kluyveromyc	792	38	39.6	222	2	Q9RUP0_DEIRA	Q9RUP0 deinococcus	
720	38.5	40.1	475	2	Q5L562_CHLAB	Q5L562 chlamydomyc	793	38	39.6	224	2	Q63V33_BURPS	Q63V33 burkholderi	
721	38.5	40.1	485	2	Q5U3X7_RAT	Q5U3X7 rattus norv	794	38	39.6	225	2	Q9FOB4_ZEAZE	Q9FOB4 zea mays (m	
722	38.5	40.1	488	2	Q30562_CAUCR	Q30562 caulobacter	795	38	39.6	230	2	Q7SDX2_NEUCR	Q7SDX2 neurospora	
723	38.5	40.1	488	2	Q9A389_CAUCR	Q9A389 caulobacter	796	38	39.6	234	2	Q4MZV8_THEPA	Q4MZV8 theileria p	
724	38.5	40.1	500	2	Q4S426_TETNG	Q4S426 tetraodon n	797	38	39.6	238	2	Q4FTI7_9GAMM	Q4FTI7 psychrobact	
725	38.5	40.1	503	2	Q8HZK1_CANFA	Q8HZK1 canis fami	798	38	39.6	239	2	Q8H312_ORYSA	Q8H312 oryza sativ	
726	38.5	40.1	504	1	ICL_MOUSE	Q92790 mus musculu	799	38	39.6	243	2	Q87GH4_VIBPA	Q87GH4 vibrio para	
727	38.5	40.1	570	1	HAO_NITEU	Q50925 nitrosomona	800	38	39.6	246	2	Q4HBUT_9DEIO	Q4HBUT deinococcus	
728	38.5	40.1	570	2	Q9RAH7_9PROT	Q9RAH7 nitrosomona	801	38	39.6	248	2	Q75EC4_ASHGO	Q75EC4 ashbya goss	
729	38.5	40.1	570	2	Q9RAH8_9PROT	Q9RAH8 nitrosomona	802	38	39.6	261	2	Q5NMN5_ZYMMO	Q5NMN5 zymomonas m	
730	38.5	40.1	570	2	Q9RAH9_9PROT	Q9RAH9 nitrosomona	803	38	39.6	268	2	Q9Y0B0_RALSO	Q9Y0B0 ralsconia s	
731	38.5	40.1	577	2	Q54214_DICDI	Q54214 dictyosteli	804	38	39.6	268	2	Q5YX88_NOCFA	Q5YX88 nocardia fa	
732	38.5	40.1	825	2	Q74ZNO_ASHGO	Q74ZNO ashbya goss	805	38	39.6	272	2	Q4NY11_9DELT	Q4NY11 anaeromyxob	
733	38.5	40.1	1299	2	Q8TXZ8_METKA	Q8TXZ8 methanopyru	806	38	39.6	274	2	Q4Q2C9_LEIMA	Q4Q2C9 leishmania	
734	38.5	40.1	1338	2	Q4NSG3_9DELT	Q4NSG3 anaeromyxob	807	38	39.6	277	2	Q761J8_PRUPB	Q761J8 prunus pers	
735	38.5	40.1	1341	2	Q4PI63_USTWA	Q4PI63 ustilago ma	808	38	39.6	279	2	Q6DW18_9GAMM	Q6DW18 acinetobact	
736	38.5	40.1	1485	2	Q85153_PHOLU	Q85153 photorhabdu	809	38	39.6	280	2	Q8NA90_HUMAN	Q8NA90 homo sapien	
737	38.5	40.1	1485	2	Q7N936_PHOLL	Q7N936 photorhabdu	810	38	39.6	280	2	Q41602_TRITU	Q41602 triticum tu	
738	38.5	40.1	1573	2	Q4IDX2_GIBZE	Q4IDX2 gibberella	811	38	39.6	281	2	Q5JGR8_PYRKO	Q5JGR8 pyrococcus	
739	38.5	40.1	1644	2	Q8Z4N7_SALTI	Q8Z4N7 salmonella	812	38	39.6	281	2	Q7NII5_GLOVI	Q7NII5 gloebacter	
740	38.5	40.1	1644	2	Q8ZN46_SALTY	Q8ZN46 salmonella	813	38	39.6	282	2	Q84M19_TRITU	Q84M19 triticum tu	
741	38.5	40.1	1644	2	Q5PNH4_SALPA	Q5PNH4 salmonella	814	38	39.6	283	2	Q5LM08_SILPO	Q5LM08 silicibacte	
742	38.5	40.1	1653	1	YFHM_ECOLI	P76578 escherichia	815	38	39.6	285	2	Q4Q694_WHEAT	Q4Q694 triticum ae	
743	38.5	40.1	1653	2	Q8CVT2_ECOL6	Q8CVT2 escherichia	816	38	39.6	288	2	Q8N6T2_HUMAN	Q8N6T2 homo sapien	
744	38.5	40.1	1653	2	Q8XA93_ECOS7	Q8XA93 escherichia	817	38	39.6	289	2	Q6MRD3_BDEBA	Q6MRD3 bdellovibri	
745	38.5	40.1	1653	2	Q6NEV4_CORDI	Q6NEV4 corynebacte	818	38	39.6	290	2	Q9QH58_9ALPH	Q9QH58 gallid herp	
746	38.5	40.1	2102	2	Q7TPK2_RAT	Q7TPK2 rattus norv	819	38	39.6	293	2	Q86594_STRCO	Q86594 streptomyce	
747	38	39.6	23	2	Q9S8B0_SECCCE	Q9S8B0 secale cere	820	38	39.6	294	2	Q5HYA6_HUMAN	Q5HYA6 homo sapien	
748	38	39.6	61	2	Q6CZG7_ERWCT	Q6CZG7 erwina car	821	38	39.6	296	1	CD2C_DICDI	P34112 dictyosteli	
749	38	39.6	81	2	Q9S7J4_ORYLO	Q9S7J4 oryza longi	822	38	39.6	296	2	Q558T7_DICDI	Q558T7 dictyosteli	
750	38	39.6	86	2	Q5SK21_THET8	Q5SK21 thermus the	823	38	39.6	298	2	Q9HXG1_PSEAB	Q9HXG1 pseudomonas	
751	38	39.6	86	2	Q72LD3_THET2	Q72LD3 thermus the	824	38	39.6	302	1	GDBX_WHEAT	P21292 triticum ae	
752	38	39.6	92	2	Q5BSZ3_SCHJA	Q5BSZ3 schistosoma	825	38	39.6	302	2	Q94G93_WHEAT	Q94G93 triticum ae	
753	38	39.6	92	2	Q5F5Q9_NEIG1	Q5F5Q9 neisseria g	826	38	39.6	307	2	Q9FXK7_ARATH	Q9FXK7 arabidopsis	
754	38	39.6	92	2	Q9JX01_NEIMA	Q9JX01 neisseria m	827	38	39.6	308	1	CAH6_HUMAN	P23280 homo sapien	
755	38	39.6	92	2	Q9K1J4_NEIME	Q9K1J4 neisseria m	828	38	39.6	312	2	Q6NSE0_RHOPA	Q6NSE0 rhodopseudo	
756	38	39.6	94	2	Q5P0K5_AZOSE	Q5P0K5 azoarcus sp	829	38	39.6	312	2	Q6ACP4_LEIXX	Q6ACP4 leifsonia x	
757	38	39.6	95	2	Q5KRL4_CORGL	Q5KRL4 corynebacte	830	38	39.6	313	2	Q5FC00_HUMAN	Q5FC00 homo sapien	
758	38	39.6	99	2	Q9Z497_9ZZZZ	Q9Z497 plasmid col	831	38	39.6	316	2	Q81606_SORBI	Q81606 sorghum bic	
759	38	39.6	99	2	Q4FIF2_GENTR	Q4FIF2 salmonella	832	38	39.6	324	2	Q82E11_STRAW	Q82E11 streptomyce	
760	38	39.6	104	2	Q8Y319_LISMO	Q8Y319 listeria mo	833	38	39.6	326	2	Q6AGX2_LEIXX	Q6AGX2 leifsonia x	
761	38	39.6	104	2	Q926R3_LISIN	Q926R3 listeria in	834	38	39.6	329	1	IPNS_STRCL	P10621 streptomyce	



835	38	39.6	329	2	Q8H958_MARPO	Q8h958_marchantia	908	38	39.6	518	2	P95535_ALCEU	P95535_alcaligenes
836	38	39.6	333	2	Q4ZRL8_PSESM	Q4zrl8_pseudomonas	909	38	39.6	518	2	Q5GRB0_ALCXX	Q5grb0_alcaligenes
837	38	39.6	333	2	Q87zt2_PSESM	Q87zt2_pseudomonas	910	38	39.6	518	2	Q8GFH4_ALCXX	Q8gfH4_alcaligenes
838	38	39.6	339	2	Q6CSM3_YARL1	Q6csm3_yarrowia li	911	38	39.6	525	2	Q5ASB0_EMENI	Q5asb0_aspergillus
839	38	39.6	339	2	Q01608_CABEL	Q01608_caenorhabdi	912	38	39.6	525	2	Q44302_ANASP	Q44302_anabaena ep
840	38	39.6	340	2	Q8GWL4_ARATH	Q8gwl4_arabidopsis	913	38	39.6	529	2	Q4RCN9_TETNG	Q4rcn9_tetradodon n
841	38	39.6	342	2	Q5ZEM9_ORYSA	Q5zem9_oryza sativ	914	38	39.6	530	1	F2B_EMENI	F48457_emeritella
842	38	39.6	351	2	Q8ZWM6_BRAJA	Q8zwm6_bradyrhizob	915	38	39.6	532	2	Q8KOH1_MOUSE	Q8koh1_mus musculus
843	38	39.6	351	2	Q58EG5_BRARE	Q58eg5_brachydanio	916	38	39.6	534	1	YDHS_ECOLI	Y77148_escherichia
844	38	39.6	356	2	Q4KGQ1_PSEPF5	Q4kgq1_pseudomonas	917	38	39.6	534	1	YDHS_ECOLI	Y77148_escherichia
845	38	39.6	357	1	LD0X_MALDO	P51091_malus domes	918	38	39.6	534	2	Q8X619_ECOS7	Q8x619_echerichia
846	38	39.6	357	2	Q8L5N4_MALDO	Q8l5n4_malus domes	919	38	39.6	536	2	Q55NT7_CRYNE	Q55nt7_cryptococcus
847	38	39.6	357	2	Q5MZN6_SYNF6	Q5mzn6_synecococc	920	38	39.6	536	2	Q4H2B4_PHACH	Q4h2b4_phanerocha
848	38	39.6	357	2	Q6NBP1_RHOPA	Q6nbp1_rhodopseudo	921	38	39.6	536	2	Q8FH66_ECOL6	Q8fh66_echerichia
849	38	39.6	373	2	Q9LDD5_ARATH	Q9ldd5_arabidopsis	922	38	39.6	541	2	Q67G34_9ACTO	Q67g34_streptomyce
850	38	39.6	374	2	Q71223_LISMF	Q71223_listeria mo	923	38	39.6	552	2	Q732Y8_MYCPA	Q732y8_mycobacteri
851	38	39.6	374	2	Q8Y6N8_LISMF	Q8y6n8_listeria mo	924	38	39.6	552	2	Q93AG2_PSEPU	Q93ag2_pseudomonas
852	38	39.6	374	2	Q92B63_LISIN	Q92b63_listeria in	925	38	39.6	554	2	Q56V45_ARATH	Q56v45_arabidopsis
853	38	39.6	376	1	P53_ICTFU	Q93379_ictalurus p	926	38	39.6	554	2	Q5EC55_MOUSE	Q5ec55_mus musculus
854	38	39.6	376	2	Q6AAZ8_PROAC	Q6aa28_propionibac	927	38	39.6	555	1	MIS_MOUSE	P27106_mus musculus
855	38	39.6	376	2	Q92KW1_RHIME	Q92kw1_rhizobium m	928	38	39.6	559	2	Q6BKL1_DEBHA	Q6bkl1_debaryomyce
856	38	39.6	376	2	Q9Q8J4_9POXV	Q9q8j4_myxoma viru	929	38	39.6	561	2	Q61087_TRYCR	Q61087_trypanosoma
857	38	39.6	376	2	Q9Q8X2_9POXV	Q9q8x2_rabbit fibr	930	38	39.6	563	2	Q63Y36_BURPS	Q63y36_burkholderi
858	38	39.6	377	2	Q73TU9_MYCPA	Q73tu9_mycobacteri	931	38	39.6	569	2	Q98PW7_MYCPU	Q98pw7_mycoplasma
859	38	39.6	383	2	Q5UL08_FRAAN	Q5ul08_fragaria an	932	38	39.6	570	2	Q61ZW8_CAEBR	Q61zw8_caenorhabdi
860	38	39.6	383	2	Q5UL09_FRAAN	Q5ul09_fragaria an	933	38	39.6	571	1	IF2_THET8	P48515_thermus the
861	38	39.6	383	2	Q65792_WHEAT	Q65792_triticum ae	934	38	39.6	571	2	Q72KE8_THET2	Q72ke8_thermus the
862	38	39.6	391	2	Q5KC72_CRYNE	Q5kc72_cryptococc	935	38	39.6	577	2	Q9RYK3_DEIRA	Q9ryk3_deinococcus
863	38	39.6	391	2	Q8ROZ6_9SPHI	Q8ro26_cytophaga s	936	38	39.6	579	2	Q7SPK7_NEUCR	Q7sfk7_neurospora
864	38	39.6	391	2	Q8RQZ1_9SPHI	Q8rqz1_cytophaga s	937	38	39.6	580	2	Q6CUP2_KLULA	Q6cup2_kluyveromyc
865	38	39.6	391	2	Q8RQZ0_9SPHI	Q8rqz0_cytophaga s	938	38	39.6	589	2	Q7TFJ8_RHOBA	Q7um27_rhodospirell
866	38	39.6	391	2	Q8RQZ0_9SPHI	Q8rqz0_cytophaga s	939	38	39.6	589	2	Q7TFJ8_RHOBA	Q7um27_rhodospirell
867	38	39.6	391	2	Q8RQZ5_9SPHI	Q8rqz5_cytophaga s	940	38	39.6	593	2	Q5AYU9_EMENI	Q5ayy9_aspergillus
868	38	39.6	391	2	Q8RQZ7_9SPHI	Q8rqz7_cytophaga s	941	38	39.6	593	2	Q8NKB1_EMENI	Q8nkb1_emeritella
869	38	39.6	391	2	Q8ROY9_9SPHI	Q8roy9_cytophaga s	942	38	39.6	603	2	Q6C4M8_YARLI	Q6c4m8_yarrowia li
870	38	39.6	391	2	Q8RQZ4_9SPHI	Q8rqz4_cytophaga s	943	38	39.6	603	2	Q829H5_STRAM	Q829h5_sreptomyc
871	38	39.6	391	2	Q8RQZ6_9SPHI	Q8rqz6_cytophaga s	944	38	39.6	613	2	Q7KT44_DROME	Q7kt44_drosophila
872	38	39.6	391	2	Q8RQZ3_9SPHI	Q8rqz3_cytophaga s	945	38	39.6	617	2	Q4UF76_THEAN	Q4uf76_theliera a
873	38	39.6	391	2	Q8RQZ2_9SPHI	Q8rqz2_cytophaga s	946	38	39.6	617	2	Q9LHC3_ARATH	Q9lhc3_arabidopsis
874	38	39.6	391	2	Q8ROY7_9SPHI	Q8roy7_cytophaga s	947	38	39.6	626	2	Q4I238_AZOV1	Q4i238_azotobacter
875	38	39.6	399	2	Q655C7_ORYSA	Q655c7_oryza eativ	948	38	39.6	627	2	Q82CK4_STRAM	Q82ck4_streptomyce
876	38	39.6	401	2	Q5RKP0_MOUSE	Q5rkp0_mus musculu	949	38	39.6	635	1	SYT_CHLCV	Q82l13_chlamydophi
877	38	39.6	404	2	Q9EW14_STRCO	Q9ew14_streptomyc	950	38	39.6	635	2	Q9DIJ7_9HIV1	Q9dij7_human immun
878	38	39.6	406	1	VAT1_MOUSE	Q62465_mus musculu	951	38	39.6	639	2	Q4MU59_BACCE	Q4mu59_bacillus ce
879	38	39.6	407	2	Q9VSH6_DROME	Q9vsh6_drosophila	952	38	39.6	640	2	Q738H6_BACCL	Q738h6_bacillus ce
880	38	39.6	413	2	Q9AND6_BRAJA	Q9and6_bradyrhizob	953	38	39.6	640	2	Q81DN5_BACCR	Q81dn5_bacillus ce
881	38	39.6	416	2	Q4IU74_AZOV1	Q4iu74_azotobacter	954	38	39.6	644	1	CR3AA_BACTD	Q82ck4_streptomyce
882	38	39.6	426	2	Q92NP1_RHIME	Q92np1_rhizobium m	955	38	39.6	644	1	CR3AA_BACTM	Q82l13_chlamydophi
883	38	39.6	432	2	Q9CA41_ARATH	Q9ca41_arabidopsis	956	38	39.6	644	1	CR3AA_BACTT	Q9dij7_human immun
884	38	39.6	438	2	Q9HP54_HALSA	Q9hp54_halobacteri	957	38	39.6	644	2	Q6FIT9_MESFL	Q6fit9_mesoplasma
885	38	39.6	440	2	Q8OW26_MOUSE	Q8ow26_mus musculu	958	38	39.6	650	2	Q4WMP6_ASFPF	Q4wmp6_aspergillus
886	38	39.6	445	2	Q9RTZ3_DEIRA	Q9rtz3_deinococcus	959	38	39.6	652	2	Q6PXN8_BACTU	Q6pxn8_bacillus th
887	38	39.6	450	2	Q8XVNO_FALSO	Q8xvno_ralstonia s	960	38	39.6	652	2	Q643Z7_BACTU	Q643z7_bacillus th
888	38	39.6	451	2	Q93SS8_FLESH	Q93ss8_plesiomonas	961	38	39.6	652	2	Q9S6N9_BACTU	Q9s6n9_bacillus th
889	38	39.6	468	2	Q73W17_MYCPA	Q73w17_mycobacteri	962	38	39.6	663	2	Q4P788_USRNA	Q4p788_ustiliago ma
890	38	39.6	468	2	Q6N2N3_RHOPA	Q6n2n3_rhodopseudo	963	38	39.6	664	2	Q4WGH7_ASFPF	Q4wgh7_aspergillus
891	38	39.6	469	1	Y1571_AERPE	Q9ybwt7_aeropyrum p	964	38	39.6	677	2	Q9FIM7_ARATH	Q9fim7_arabidopsis
892	38	39.6	470	2	Q7UWJ4_RHOBA	Q7uwj4_rhodospirell	965	38	39.6	684	1	KBTE7_HUMAN	Q8wvz9_homo sapien
893	38	39.6	474	2	Q9AJH6_9SPHI	Q9ajh6_cytophaga s	966	38	39.6	684	2	Q6F168_HUMAN	Q6f168_homo sapien
894	38	39.6	476	2	Q6CSV2_YARLI	Q6c6v2_yarrowia li	967	38	39.6	684	2	Q5T6Y7_HUMAN	Q5t6y7_homo sapien
895	38	39.6	478	1	BMP3B_HUMAN	P55107_homo sapien	968	38	39.6	684	2	Q9VUJ39_DROME	Q9vj39_drosophila
896	38	39.6	478	2	Q5VSQ8_HORPE	Q5vsg8_homo sapien	969	38	39.6	684	2	Q63J44_BURPS	Q63j44_burkholderi
897	38	39.6	478	2	Q7WVP3_BORPE	Q7wvp3_bordetella	970	38	39.6	684	2	Q62E33_BURMA	Q62e33_burkholderi
898	38	39.6	478	2	Q7W916_BORPA	Q7w916_bordetella	971	38	39.6	685	1	IF2_CLOTE	Q89518_clostridium
899	38	39.6	478	2	Q7WH74_BORER	Q7wh74_bordetella	972	38	39.6	686	2	Q4NZP5_9DELT	Q4nzp5_aeromonas
900	38	39.6	480	2	Q6TKY7_ECOLI	Q6tk7_echerichia	973	38	39.6	688	2	Q8KU83_ENTFA	Q8ku83_enterococc
901	38	39.6	492	2	Q6AEN7_LEIXX	Q6aen7_leifsonia x	974	38	39.6	690	2	Q501K2_MOUSE	Q501k2_mus musculu
902	38	39.6	496	2	Q01863_PHYCP	Q01863_phytophthor	975	38	39.6	710	2	Q9FIM6_ARATH	Q9fim6_arabidopsis
903	38	39.6	499	2	Q9P713_NEUCR	Q9p713_neurospora	976	38	39.6	713	2	Q9FIM5_ARATH	Q9fim5_arabidopsis
904	38	39.6	502	2	Q88TS9_LACPL	Q88ts9_lactobacill	977	38	39.6	723	2	Q6XVH4_LACSK	Q6xvh4_lactobacill
905	38	39.6	506	2	Q5QSR5_TRIVA	Q5qsr5_trichomonas	978	38	39.6	730	1	PLOD_CAEEL	Q20679_caenorhabdi
906	38	39.6	506	2	Q4NEZ2_9MICC	Q4nez2_arthrobacte	979	38	39.6	732	1	BOPI_MOUSE	P97452_mus musculu
907	38	39.6	513	1	YWFF_LACLA	Q9cdp0_lactococcus	980	38	39.6	734	2	Q89RK9_BRAJA	Q89rk9_bradyrhizob

981 38 39.6 737 2 Q6ZQ19 mus musculus  
 982 38 39.6 743 2 O58900 pyrococcus  
 983 38 39.6 750 1 MUTA\_PIG  
 984 38 39.6 754 2 Q4RX33 TETNG  
 985 38 39.6 757 1 CEP68 HUMAN  
 986 38 39.6 765 2 Q61UW4 CAEBR  
 987 38 39.6 767 2 Q6BX13 DEBHA  
 988 38 39.6 771 2 Q9BQ17 HUMAN  
 989 38 39.6 771 2 Q7NVY6 CHRVO  
 990 38 39.6 775 1 DPOL\_THES9  
 991 38 39.6 776 2 Q7NVY1 PHOLL  
 992 38 39.6 785 2 Q5B2R9 EMENT  
 993 38 39.6 792 2 Q9H0D2 HUMAN  
 994 38 39.6 798 2 Q42698 ASPOR  
 995 38 39.6 798 2 O59862 ASPOR  
 996 38 39.6 824 2 Q8U7P1 AGRT5  
 997 38 39.6 828 1 NAPA\_ECOLI  
 998 38 39.6 828 2 Q65503 ARATH  
 999 38 39.6 828 2 Q9CKL8 PASMU  
 1000 38 39.6 834 2 Q7CUV7\_AGR5

## ALIGNMENTS

RESULT 1  
 TDGF1 HUMAN STANDARD; PRT; 188 AA.  
 AC P13385;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Teratocarcinoma-derived growth factor 1 precursor (Epidermal growth factor-like growth factor) (CRGF).  
 DE factor-like growth factor 1 precursor (CRGF).  
 GN Name=TDGF1; Synonyms=CRPTO;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP MEDLINE=90005403; PubMed=2792079;  
 RA Ciccodicola A., Dono R., Obici S., Zollo M., Persico M.G.;  
 RT "Molecular characterization of a gene of the 'EGF family' expressed in undifferentiated human NTERA2 teratocarcinoma cells.";  
 RL EMBO J. 8:1987-1991(1989).  
 RN [2]  
 RP MEDLINE=91353571; PubMed=1882841;  
 RA Dono R., Montuori N., Rocchi M., de Ponti-Zilli L., Ciccodicola A., Persico M.G.;  
 RT "Isolation and characterization of the CRPTO autosomal gene and its X-linked related sequence.";  
 RL Am. J. Hum. Genet. 49:555-565(1991).  
 RN [3]  
 RP PROTEIN SEQUENCE OF 31-45.  
 RX PubMed=15340161; DOI=10.1101/ps.04682504;  
 RA Zhang Z., Henzel W.J.;  
 RT "Signal peptide prediction based on analysis of experimentally verified cleavage sites.";  
 RL Protein Sci. 13:2819-2824(2004).  
 RN [4]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=22800780; PubMed=12919325;  
 RX DOI=10.1046/j.1432-1033.2003.03749.x;  
 RA Foley S.F., Van Vlijmen H.W., Boynton R.E., Cheung A.E., Singh J., Sanicola M., Young C.N., Wen D.;  
 RT "The CRPTO/FRL-1/CRYPIC (CFC) domain of human Crpto.";  
 RL Eur. J. Biochem. 270:3610-3618(2003).  
 CC -!- FUNCTION: Could play a role in the determination of the epiblastic cells that subsequently give rise to the mesoderm.  
 CC -!- TISSUE SPECIFICITY: Preferentially expressed in gastric and

colorectal carcinomas than in their normal counterparts.  
 -!- SIMILARITY: Contains 1 EGF-like domain.  
 -----  
 This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
 -----  
 CC EMBL; M96955; AAA61134.1; -; Genomic\_DNA.  
 CC EMBL; X14253; CAA32467.1; -; mRNA.  
 CC PIR; B39787; A30362.  
 CC HSPP; P00740; IEDM.  
 CC Ensembl; ENSG00000163828; Homo sapiens.  
 CC HGNC; HGNC:11701; TDGF1.  
 CC H-invDB; Hix0003250; -.  
 CC MIM; 187395; -.  
 CC GO; GO:0009986; C:cell surface; IDA.  
 CC GO; GO:0019897; C:extrinsic to plasma membrane; ISS.  
 CC GO; GO:0008083; F:growth factor activity; IDA.  
 CC GO; GO:000187; P:activation of MAPK; IDA.  
 CC GO; GO:0001763; P:branching morphogenesis; TAS.  
 CC GO; GO:0030154; P:cell differentiation; TAS.  
 CC GO; GO:0008595; P:determination of anterior/posterior axis, e. . .; ISS.  
 CC GO; GO:0009790; P:embryonic development; TAS.  
 CC GO; GO:0007507; P:heart development; IDA.  
 CC GO; GO:0030879; P:mammary gland development; TAS.  
 CC GO; GO:0043066; P:negative regulation of apoptosis; IDA.  
 CC GO; GO:0018105; P:peptidyl-serine phosphorylation; IDA.  
 CC GO; GO:0030335; P:positive regulation of cell migration; IDA.  
 CC GO; GO:0008284; P:positive regulation of cell proliferation; TAS.  
 CC GO; GO:0050731; P:positive regulation of peptidyl-tyrosine ph. . .; IDA.  
 CC GO; GO:0009966; P:regulation of signal transduction; IDA.  
 CC InterPro; IPR000742; EGF\_2.  
 CC InterPro; IPR006209; EGF\_like.  
 CC InterPro; IPR006210; IEGF.  
 CC Pfam; PF00008; EGF; 1.  
 CC SMART; SM00181; EGF; 1.  
 CC PROSITE; PS00022; EGF\_1; 1.  
 CC PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 CC PROSITE; PS00026; EGF\_3; 1.  
 CC Direct protein sequencing; EGF-like domain; Glycoprotein;  
 KW Growth factor; Polymorphism; Signal.  
 FT SIGNAL 1 30  
 FT CHAIN 31 188 Teratocarcinoma-derived growth factor 1.  
 FT DOMAIN 78 107 EGF-like. (GlcNAc. . .) (Potential).  
 FT CARBOHYD 79 79 N-linked. (GlcNAc. . .) (Potential).  
 FT DISULFID 82 89 By similarity.  
 FT DISULFID 83 95 By similarity.  
 FT DISULFID 97 106 By similarity.  
 FT DISULFID 115 133 By similarity.  
 FT DISULFID 128 149 By similarity.  
 FT DISULFID 131 140 By similarity.  
 FT VARIANT 43 43 Y -> D (in dbSNP:2293025).  
 FT  
 SQ SEQUENCE 188 AA; 21169 MW; ABE8727D0F27D886 CRC64;  
 Query Match 100.0%; Score 96; DB 1; Length 188;  
 Best Local Similarity 100.0%; Pred. NO. 6.7e-07;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FRDSDIWPQEEPAIRPR 17  
 |||||  
 DB 46 FRDSDIWPQEEPAIRPR 62  
 RESULT 2  
 TDGF2 HUMAN STANDARD; PRT; 188 AA.  
 ID TDGF2\_HUMAN  
 AC P51864;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)



DE Teratocarcinoma-derived growth factor 2 (Epidermal growth factor-like  
 GN Cripito protein CR3) (Cripito-3 growth factor).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Lung fibroblast;  
 RX MEDLINE=91353571; PubMed=1882841;  
 RA Dono R., Montuori N., Rocchi M., de Ponti-Zilli L., Ciccodicola A.,  
 RA Persico M.G.;  
 RT "Isolation and characterization of the CRYPTO autosomal gene and its  
 RT X-linked related sequence."  
 RL Am. J. Hum. Genet. 49:555-565(1991).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Johnson D., Wamsley P., Gibson A.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Kintner M.A., Kintner R.L., Hosick H.L.;  
 RT "Cripito-3 and Cripito-1 have different effects on the growth  
 RT characteristics of MCF-7 and Vero cells."  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Could play a role in the determination of the epiblastic  
 CC cells that subsequently give rise to the mesoderm.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR ENBL; M96956; AAA61135.1; -; mRNA.  
 DR ENBL; AC000113; AAB46353.1; -; Genomic DNA.  
 DR ENBL; AF251549; AAG49538.1; -; Genomic DNA.  
 DR ENBL; AF251550; AAG49539.1; -; Genomic DNA.  
 DR HSSP; P00740; 1EDM  
 DR Ensembl; ENSG00000183514; Homo sapiens.  
 DR HGNC; HGNC:11703; TDGF3.  
 DR MIM; 187395; -;  
 DR GO; GO:0008083; F: growth factor activity; NAS.  
 DR GO; GO:0007500; P: mesodermal cell fate determination; NAS.  
 DR InterPro; IPR000742; EGF\_2.  
 DR InterPro; IPR006209; EGF\_like.  
 DR Pfam; PF00008; EGF\_1.  
 DR PROSITE; PS00022; EGF 1; 1.  
 DR PROSITE; PS00026; EGF\_3; FALSE NEG.  
 KW EGF-like domain; Glycoprotein; Growth factor.  
 FT DOMAIN 78 107 EGF-like.  
 FT CARBOHYD 79 79 N-linked (GlcNAc...) (Potential).  
 FT DISULFID 82 89 By similarity.  
 FT DISULFID 83 95 By similarity.  
 FT DISULFID 97 106 By similarity.  
 FT DISULFID 115 133 By similarity.  
 FT DISULFID 128 149 By similarity.  
 FT DISULFID 131 140 By similarity.  
 SQ SEQUENCE 188 AA; 21181 MW; C07AC973E4D82B32 CRC64;

Query Match 100.0%; Score 96; DB 1; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-07; Indels 0; Gaps 0;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FRDSDIWPQEEPAIRPR 17  
 |||||  
 DB 46 FRDSDIWPQEEPAIRPR 62

RESULT 3

Q8TCC1\_HUMAN  
 ID Q8TCC1\_HUMAN PRELIMINARY; PRT; 188 AA.  
 AC Q8TCC1;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Teratocarcinoma-derived growth factor 1.  
 GN Names=TDGF1;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain, and Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain;  
 RG NIH MGC Project;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR ENBL; BC022393; AAH23393.1; -; mRNA.  
 DR ENBL; BC067844; AAH67844.1; -; mRNA.  
 DR PIR; A39787; A39787.  
 DR HSSP; P00740; 1EDM.  
 DR Ensembl; ENSG00000163828; Homo sapiens.  
 DR InterPro; IPR000742; EGF\_2.  
 DR InterPro; IPR006209; EGF\_like.  
 DR Pfam; PF00008; EGF\_1.  
 DR SMART; SM00181; EGF; 1.  
 DR PROSITE; PS00022; EGF 1; UNKNOWN\_1.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 KW EGF-like domain; Transmembrane.  
 SQ SEQUENCE 188 AA; 21141 MW; E1B3A8461C855FFF CRC64;

Query Match 100.0%; Score 96; DB 2; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-07; Indels 0; Gaps 0;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FRDSDIWPQEEPAIRPR 17  
 |||||  
 DB 46 FRDSDIWPQEEPAIRPR 62

RESULT 4  
Q58D57\_BOVIN

```

ID Q58D57 BOVIN PRELIMINARY; PRT; 181 AA.
AC Q58D57
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE Teratocarcinoma-derived growth factor 1.
GN Name=TDGF1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC
RX MEDLINE=21180013; PubMed=11282978; DOI=10.1101/gr.170101;
RA Smith T.P.L., Grosse W.M., Freking B.A., Roberts A.J., Stone R.T.,
RA Casas E., Wray J.E., White J., Cho J., Fahrenkrug S.C., Bennett G.L.,
RA Heaton M.P., Laegreid W.W., Rohrer G.A., Chitko-McKown C.G.,
RA Heaton G., Holt I., Karamycheva S., Liang F., Quackenbush J.,
RA Keele J.W.;
RT "Sequence evaluation of four pooled-tissue normalized bovine cDNA
RT libraries and construction of a gene index for cattle.";
RN Genome Res. 11:626-630(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC
RX Harhay G.P., Sonstegard T.S., Van Tassel C.P., Clawson M.L.,
RA Heaton M.P., Keele J.W., Snelling W.M., Weidmann R.T., Smith T.P.L.;
RT "Sequencing and analysis of Bos taurus full-length insert cDNA
RT clones.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT021740; AA046587.1; -; mRNA.
SQ SEQUENCE 181 AA; 2023 MW; A4E8AEA78E106514 CRC64;

Query Match 72.9%; Score 70; DB 2; Length 181;
Best Local Similarity 80.08; Pred. No. 0.0081;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FRDSDIWQEPAPR 15
Db 41 FRDDGLWSQEPAPR 55

RESULT 5
TDGF1 MOUSE STANDARD; PRT; 171 AA.
AC P51865;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE Teratocarcinoma-derived growth factor precursor (Epidermal growth
DE factor-like Cripito protein) (Cripito growth factor).
GN Name=tdgf1; Synonyms=Cripito;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=C57BL/6; TISSUE=Embryo;
RX MEDLINE=94094736; PubMed=7916676;
RA Dono R., Scalera L., Pacifico F., Acampora D., Persico M.G.,
RA Simeone A.;
RT "The murine cripito gene: expression during mesoderm induction and
RT early heart morphogenesis.";
FT Development 118:1157-1168(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 135-171.
RC STRAIN=129/SV;
RX MEDLINE=96269599; PubMed=8661720; DOI=10.1007/s003359900100;
RA Liguori G., Tucci M., Montuori N., Dono R., Lago C.T., Pacifico A.F.,

```

```

RA Persico M.G.;
RT "Characterization of the mouse Tdgfl gene and Tdgd pseudogenes.";
RL Mamm. Genome 7:344-348(1996).
CC -1- FUNCTION: Could play a role in the determination of the epiblastic
CC cells that subsequently give rise to the mesoderm.
CC -1- TISSUE SPECIFICITY: Expressed at low level in specific organs of
CC the adult animal such as spleen, heart, lung and brain. During
CC gastrulation, expressed in the forming mesoderm. In later stages
CC of the developing heart, expression is restricted to the truncus
CC arteriosus.
CC -1- DEVELOPMENTAL STAGE: First expressed prior to the onset of
CC gastrulation (early streak stage), then continues throughout
CC embryonic development.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
DR EMBL; M87321; AAA37459.1; -; mRNA.
DR EMBL; X94083; CAA63827.1; -; Genomic_DNA.
DR PIR; I49612; I49612.
DR HSPSP; P00740; IEDM.
DR Ensembl; ENSMUSG00000032494; Mus musculus.
DR MGI; MGI:98658; Tdgfl.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0019897; C:extrinsic to plasma membrane; IDA.
DR GO; GO:0001763; P:branching morphogenesis; IDA.
DR GO; GO:0030154; P:cell differentiation; TAS.
DR GO; GO:0008595; P:determination of anterior/posterior axis, e. . .; IMP.
DR GO; GO:0007507; P:heart development; IDA.
DR GO; GO:0030335; P:positive regulation of cell migration; IDA.
DR InterPro; IPR000742; EGF_2.
DR Pfam; PF00008; EGF_1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00026; EGF_3; 1.
KW EGF-like domain; Glycoprotein; Growth factor; Signal.
FT SIGNAL 1 17 Potential.
FT CHAIN 18 171 Teratocarcinoma-derived growth factor.
FT DOMAIN 62 91 EGF-like.
FT CARBOHYD 63 63 N-linked (GlcNAc. . .) (Potential).
FT DISULFID 66 73 By similarity.
FT DISULFID 67 79 By similarity.
FT DISULFID 81 90 By similarity.
FT DISULFID 99 117 By similarity.
FT DISULFID 112 133 By similarity.
FT DISULFID 115 124 By similarity.
SQ SEQUENCE 171 AA; 18754 MW; C52051AEACDB5380 CRC64;

Query Match 66.7%; Score 64; DB 1; Length 171;
Best Local Similarity 68.8%; Pred. No. 0.067;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RDSINWQEPAPR 17
Db 31 RDSINWQEPAPR 46

RESULT 6
Q7TQ06 MOUSE PRELIMINARY; PRT; 171 AA.
ID Q7TQ06
AC Q7TQ06;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 25, Last annotation update)
DE Tdgfl protein.
DE Name=tdgf1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

```



DR InterPro: IPR004584; Rad50.  
 DR TIGRFAMs: TIGR00606; rad50; 1.  
 KW Repeat; Transport.  
 SQ SEQUENCE 1312 AA; 151236 MW; 550910E426F9106B CRC64;

Query Match 51.0%; Score 49; DB 2; Length 1312;  
 Best Local Similarity 46.2%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 RDSIWPQEEPPI 14  
 DB 155 QDESILWPMSEPSV 167  
 :|:|:| |::|

RESULT 9  
 Q5B761\_EMENI PRELIMINARY; PRT; 1319 AA.  
 ID AC Q5B761;  
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=AN3619.2;  
 OS Aspergillus nidulans FGSC A4.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiiales; Trichocomaceae; Emericella.  
 OX NCBI\_TaxID=227321;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=FGSC A4;  
 RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,  
 RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,  
 RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,  
 RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,  
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,  
 RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,  
 RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,  
 RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,  
 RA Jaffe D., Johnson R., Jones C., Kamal M., Karatas A.,  
 RA Kells S., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,  
 RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,  
 RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,  
 RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,  
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,  
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,  
 RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,  
 RA Roman J., Schauer S., Schupbach R., Seaman S., Severy P., Smirnov S.,  
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,  
 RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,  
 RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,  
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,  
 RA Lander E.;  
 RT "Genome Sequence of Aspergillus nidulans";  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AACD0100061; EAA59827.1; -; Genomic\_DNA.  
 DR GO; GO:0030870; C:Mre11 complex; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0016887; F:ATPase activity; IEA.  
 DR GO; GO:0006281; P:DNA repair; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003439; ABC transp\_like.  
 DR InterPro; IPR004584; Rad50.  
 DR TIGRFAMs; TIGR00606; rad50; 1.  
 KW Hypothetical protein; Repeat; Transport.  
 SQ SEQUENCE 1319 AA; 151992 MW; 83FDAFB7C96BE36 CRC64;

Query Match 51.0%; Score 49; DB 2; Length 1319;  
 Best Local Similarity 46.2%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 RDSIWPQEEPPI 14  
 Db 174 QDESILWPMSEPSV 186  
 :|:|:| |::|

```

RESULT 12
Q22352 CAEEL PRELIMINARY; PRT; 333 AA.
AC Q22352;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1999 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein T08H10.1.
GN ORFNames=T08H10.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBTaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT Investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; U55368; AAA97992.2; -; Genomic_DNA.
DR PIR; B89027; B90027.
DR HSP; P14550; 2ALR.
DR Ensembl; T08H10.1; Caenorhabditis elegans.
DR WormBase; WBGene00020369; T08H10.1.
DR WormPep; T08H10.1; CE17230.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR001395; Aldo/ket_red.
DR Pfam; PF00248; Aldo_ket_red; 1.
DR PRINTS; PR00069; ALDKETREDTASE.
DR PRODOM; PD000288; Aldo/ket_red; 1.
DR PROSITE; PS00798; ALDO-KETO REDUCTASE 1; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 333 AA; 37315 MW; 54D0333B945183B09 CRC64;

Query Match 50.0%; Score 48; DB 2; Length 333;
Best Local Similarity 40.0%; Pred. No. 47;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 RDSIWQSEPAIRP 16
Db 223 RPDGVWPEGDPLEP 237

RESULT 13
Q7KVA9 DROME
AC Q7KVA9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CG7955-PC, isoform C (GH20617P).
GN Names=CG7955; ORFNames=CG7955;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBTaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.G., Morten J.R., Richards S., Ashburner M., Henderson S.N.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballwe R.M., Basu A., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Beeson K.Y., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

```

Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
[7]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=Berkley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
Yu C., Lewis S.E., Rubin G.M., Celnikier S.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to the ABC transporter family.  
DR EMBL; AE003472; AAF47526.1; -; Genomic\_DNA.  
DR ENBL; AY051556; AAK92980.1; -; mRNA.  
DR Ensemble; CG7955; Drosophila melanogaster.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0005524; F: ATP binding; IEA.  
DR GO; GO:0016887; F: ATPase activity; IEA.  
DR GO; GO:0042626; F: ATPase activity, coupled to transmembrane m. .; IEA.  
DR GO; GO:0000166; F: nucleotide binding; IEA.  
DR GO; GO:0006810; P: transport; IEA.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR011527; ABC\_membrane\_1.  
DR InterPro; IPR01140; ABC\_TM\_transp.  
DR InterPro; IPR003439; ABC\_transp\_like.  
DR Pfam; PF00664; ABC\_membrane; 1.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR ProDom; PD000006; ABC\_transporter; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS00929; ABC\_TM1P; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 1.  
KW ATP-binding; Nucleotide-binding; Repeat; Transport.  
SQ SEQUENCE 606 AA; 66104 MW; 76C80500A1B62327 CRC64;  
  
Query Match 50.0%; Score 48; DB 2; Length 606;  
Best Local Similarity 58.3%; Pred. No. 90;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
QY 6 IWQEEPAIRPR 17  
Db |||:|:|:|  
9 IWPKEDPLVRKR 20  
  
RESULT 14  
Q7KVB1\_DROME PRELIMINARY; PRT; 709 AA.  
AC Q7KVB1;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE CG7955-PB, isoform B.  
GN Name=CG7955; ORFNames=CG7955;  
OS Drosophila melanogaster (fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN NUCLEOTIDE SEQUENCE.  
[1]  
RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
R Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Posler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.H., Wang X.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhu S., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RA "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
[2]  
RN NUCLEOTIDE SEQUENCE.  
RP MEDLINE=22426065; PubMed=12537568;  
RA Celnikier S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
melanogaster euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
[3]  
RN NUCLEOTIDE SEQUENCE.  
RP MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.N., Kronmiller B., Carlson J.W., Svirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celnikier S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
a genomics perspective.";  
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
[4]  
RN NUCLEOTIDE SEQUENCE.  
RP MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celnikier S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
systematic review.";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
[5]  
RN NUCLEOTIDE SEQUENCE.  
RP Berkeley Drosophila Genome Project;  
RA Celnikier S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,  
RA Yu C., Rubin G.;  
RT "Drosophila melanogaster release 4 sequence.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[6]  
RN NUCLEOTIDE SEQUENCE.  
RP FlyBase;  
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to the ABC transporter family.  
DR EMBL; AE003472; AAF47524.3; -; Genomic\_DNA.



DR Ensembl; CG7955; Drosophila melanogaster.  
 DR GO: GO:0016021; C: integral to membrane; IEA.  
 DR GO: GO:0005524; F: ATP binding; IEA.  
 DR GO: GO:0016887; F: ATPase activity; IEA.  
 DR GO: GO:0042626; F: ATPase activity, coupled to transmembrane m. . . ; IEA.  
 DR GO: GO:0000166; F: nucleotide binding; IEA.  
 DR GO: GO:0006810; P: transport; IEA.  
 DR InterPro: IPR003593; AAA ATPase.  
 DR InterPro: IPR011527; ABC membrane 1.  
 DR InterPro: IPR001140; ABC TM transept.  
 DR InterPro: IPR003439; ABC transp\_like.  
 DR Pfam: PF00664; ABC membrane; 1.  
 DR Pfam: PF00005; ABC trans; 1.  
 DR ProDom: PD000006; ABC transporter; 1.  
 DR SMART: SM00382; AAA; 1.  
 DR PROSITE: PS00929; ABC TRANSPORTER 1; 1.  
 DR PROSITE: PS00211; ABC TRANSPORTER 1; 1.  
 DR PROSITE: PS00893; ABC TRANSPORTER 2; 1.  
 KW ATP-binding; Nucleotide-binding; Repeat; Transport.  
 SQ SEQUENCE 709 AA; 76743 MW; FDA88EEA033CBBF1 CRC64;  
 Query Match 50.0%; Score 48; DB 2; Length 709;  
 Best Local Similarity 58.3%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 6 IWFOEPAIRPR 17  
 |||:|:|:  
 Db 112 IMPKEDPLVRKR 123  
 RESULT 15  
 Q9W0C5 DROME  
 ID Q9W0C5 DROME PRELIMINARY; PRT; 743 AA.  
 AC Q9W0C5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE CG7955-PA, isoform A.  
 GN Name=CG7955; ORFNames=CG7955;  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananthides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Hewman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirkas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
 melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirkas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 a genomics perspective.";  
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 systematic review.";  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Berkeley Drosophila Genome Project;  
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
 RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirkas R., Smith E.,  
 RA Yu C., Rubin G.;  
 RT "Drosophila melanogaster release 4 sequence.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RG FlyBase;  
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
 CC -1- INTERACTION:  
 CC Q9VVE7:CG13724; NDexp=1; IntAct=EBI-143332, EBI-110098;  
 CC Q9VVT9:CG15741; NDexp=1; IntAct=EBI-143332, EBI-89697;  
 CC -1- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AB003472; AAP47525.2; -; Genomic\_DNA.  
 DR HSP; P08716; IMT0.  
 DR Insect; Q9W0C5; -;  
 DR Ensembl; CG7955; Drosophila melanogaster.  
 DR FlyBase; FBgn0035244; CG7955.  
 DR GO: GO:0016021; C: integral to membrane; IEA.  
 DR GO: GO:0005524; F: ATP binding; IEA.  
 DR GO: GO:0016887; F: ATPase activity; IEA.  
 DR GO: GO:0042626; F: ATPase activity, coupled to transmembrane m. . . ; IEA.  
 DR GO: GO:0000166; F: nucleotide binding; IEA.  
 DR GO: GO:0006810; P: transport; IEA.

```

DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS0929; ABC_TM1F; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS08993; ABC_TRANSPORTER_2; 1.
DR ATP-binding; Nucleotide-binding; Repeat; Transport.
SQ SEQUENCE 743 AA; 80467 MW; 620B7CF0F6F06A4F CRC64;

Query Match 50.0%; Score 48; DB 2; Length 743;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 6 IWQPEEPAIRPR 17
Db 146 IMPKEDPLVRKR 157

RESULT 16
KOKO YEAST
ID _KOKO YEAST STANDARD; PRT; 1081 AA.
AC Q12236;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable serine/threonine-protein kinase YOL100W (EC 2.7.1.37).
GN OrderedLocusNames=YOL100W; ORFNames=HRC1081;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96076631; PubMed=7502582;
RA Vandenbol M., Durand P., Fortetelle D., Hilger F.;
RT "Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV
RT including the Tyl-H3 retrotransposon, the sufl1(+) frameshift
RT suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-1a and a
RT delta element.";
RL Yeast 11:1069-1075(1995).
RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S288c / FY1679;
RX MEDLINE=97313270; PubMed=9169874;
RA Dujon B., Albermann K., Aldea M., Alexandraki D., Ansoorge W.,
RA Arino J., Benes V., Bohn C., Bolotin-Fukuhara M., Bordome R.,
RA Boyer J., Camasses A., Casamayor A., Casas C., Cheret G.,
RA Criepluch C., Daigman-Pornier B., Dang D.V., de Haan M., Delius H.,
RA Durand P., Fairhead C., Feldmann H., Gaillon L., Galisson F.,
RA Gamo F.-J., Gancedo C., Goffeau A., Goulding S.E., Grivell L.A.,
RA Habbig B., Hand N.-J., Hani J., Hattenhorst U., Hebling U.,
RA Hernando Y., Herrero E., Heumann K., Hiesel R., Hilger F., Hofmann B.,
RA Hollenberg C.P., Hughes B., Jauniaux J.-C., Kalogeropoulos A.,
RA Katsoulou C., Kordes E., Lafuente M.J., Landt O., Louis E.J.,
RA Maarse A.C., Madania A., Mannhaupt G., Marck C., Martin R.P.,
RA Mewes H.-W., Michaux G., Paces V., Parle-McDermott A.G., Pearson B.M.,
RA Perrin A., Petterson B., Poch O., Pohl T.M., Poirey R.,
RA Portecelle D., Fujol A., Furnelle B., Ramezani Rad M., Rechmann S.,
RA Schwager C., Schweizer M., Sor F., Sterky F., Tarassov I.A.,
RA Teodorou C., Tettelin H., Thierry A., Tobiasch E., Tzermia M.,
RA Uhlen M., Unsel M., Valens M., Vandenbol M., Vetter I., Vlack C.,
RA Voet M., Volckaert G., Voss H., Wambutt R., Wedler H., Wiemann S.,
RA Winsor B., Wolfe K.H., Zollner A., Zumstein E., Kleine K.;
RT "the nucleotide sequence of Saccharomyces cerevisiae chromosome XV.";
RL Nature 387:98-102(1997).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. PKC
CC subfamily.

```

```

CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; Z48149; CAA88162.1; -; Genomic DNA.
CC EMBL; Z74842; CAA99113.1; -; Genomic_DNA.
CC FIRM; S51899; S51899.
CC HSP; O15530; IHW.
CC IntAct; Q12236; -.
CC GeneOnline; 143522; -.
CC Ensembl; YOL100W; Saccharomyces cerevisiae.
CC SGD; S00005460; PKH2.
CC GO; GO:0005634; C:nucleus; IDA.
CC GO; GO:0004672; F:protein kinase activity; IDA.
CC GO; GO:0000196; P:MAPKKK cascade during cell wall biogenesis; IGI.
CC GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
CC InterPro; IPR000719; Prot kinase.
CC InterPro; IPR008271; Ser_thr_kin_AS.
CC InterPro; IPR002290; Ser_thr_kinase.
CC Pfam; PF00069; Pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC ATP-binding; Complete proteome; Hypothetical protein; Kinase;
CC Nucleotide-binding; Serine/threonine-protein kinase; Transferase.
FT DOMAIN 179 443
FT NP_BIND 185 193 ATP (By similarity).
FT ACT_SITE 303 303 Proton acceptor (By similarity).
FT BINDING 208 208 ATP (By similarity).
SQ SEQUENCE 1081 AA; 121660 MW; BE0DD9D49AC2EBC3 CRC64;

Query Match 50.0%; Score 48; DB 1; Length 1081;
Best Local Similarity 43.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FRDSIWQPEPAIRP 16
Db 448 FKDGVSWSKTPPEIKP 463

RESULT 17
Q6CUR2 KLJULA
ID Q6CUR2 KLJULA PRELIMINARY; PRT; 1296 AA.
AC Q6CUR2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to sp|P12753 Saccharomyces cerevisiae YNL250w RAD50 DNA repair
DE protein.
GN OrderedLocusNames=KLJA0C02915g;
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CBS 2359 / IFO 1267 / NRRL Y-1140 / WM37;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boistame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,

```



```

RESULT 19
ID Q5M455_STRT2
Q5M455_STRT2 PRELIMINARY; PRT; 248 AA.
AC Q5M455;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Nitro/flavin reductase.
OS Names=nFrA; OrderedLocusNames=stull149;
GN Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=264199;
RN [1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goltman E., Mazur M., Pusch G.D.,
RA Fonskei M., Overbeek R., Kyrides N., Purnelle B., Prozzi D.,
RA Ngui K., Masny D., Hancy F., Burtan S., Boutry M., Delcours J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
DR ENBL; CP000023; AAV60787.1; -; Genomic DNA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000415; Nitroreductase.
DR Pfam; PF00881; Nitroreductase; 1.
KW Complete proteome.
SQ SEQUENCE 248 AA; 28043 MW; BCD95ADCF3D2B24D CRC64;
Query Match 49.0%; Score 47; DB 2; Length 248;
Best Local Similarity 54.5%; Pred. No. 49;
Matches 6; Conservative 3; Mismatches 2; Indels 0;

Qy 7 WFOEPAIRPR 17
||:|:|:|
Db 163 WPDQEPQLKPR 173

RESULT 20
ID Q6NXX6_CORDI
Q6NXX6_CORDI PRELIMINARY; PRT; 254 AA.
AC Q6NXX6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative nitroreductase.
OS OrderedLocusNames=DIP0265;
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OC NCBI_TaxID=11717;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RX STRAIN=biotype gravis / NCTC 13129;
RX MEDLINE=22365543; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA Cerdano-Tarraga A.-M., Efstathiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Beers G.S., Churcher C.M., James K.D.,
RA De Zeyza A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jagsels K., Moulse S., Quail M.A.,
RA Rabinowitz E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrell B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
RT diphtheriae NCTC13129.";
RL Nucleic Acids Res. 31:6516-6523(2003).
DR ENBL; BX248354; CAE48770.1; -; Genomic DNA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000415; Nitroreductase.
DR Pfam; PF00881; Nitroreductase; 1.
KW Complete proteome.

```

```

SQ SEQUENCE 254 AA; 28091 MW; 66AEABAB10B1EEAB CRC64;
Query Match 49.0%; Score 47; DB 2; Length 254;
Best Local Similarity 54.5%; Pred.No. 50;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 7 WPOEPAIRPR 17
   ||:||||:
Db 167 WPNQEPOLKPR 177

RESULT 21
Q4NCZ5_9M1CC
ID Q4NCZ5_9M1CC PRELIMINARY; PRT; 264 AA.
AC Q4NCZ5_9M1CC
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE GN5-related N-acetyltransferase.
GN ORFNames=ArthDRAFT_0754;
OS Arthrobacter sp. FB24
OC Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
OC Micrococccineae; Micrococccaceae; Arthrobacter.
OX NCBI_TaxID=290399;
RN [1]
RC NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Arthrobacter sp. FB24."
RT Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHG01000013; EAL95237.1; -; Genomic_DNA.
KW Transferrase.
SQ SEQUENCE 264 AA; 28862 MW; 27B9E0B007AF1A57 CRC64;

Query Match 49.0%; Score 47; DB 2; Length 264;
Best Local Similarity 65.7%; Pred.No. 52;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RDSINWQEEPA 13
   |:||||:
Db 39 RANSINWPRSEPA 50

RESULT 22
Q8R0F4_MOUSE
ID Q8R0F4_MOUSE PRELIMINARY; PRT; 355 AA.
AC Q8R0F4_MOUSE
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 25, Last annotation update)
DE Igcc protein (Fragment).
GN Name=Igcc; Synonyms=BC026978;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX

```

```

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RC Director MGC Project;
RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC026978; AAH26978.1; -; mRNA.
DR Ensembl; ENSMUSG0000040795; Mus musculus.
DR MGI; MGI:2446212; BC026978.
DR MGI; MGI:2446212; Iqcc.
FT NON TER 1
SQ SEQUENCE 355 AA; 39677 MW; C2B501FC6837697D CRC64;

Query Match 49.0%; Score 47; DB 2; Length 355;
Best Local Similarity 80.0%; Pred.No. 72;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 RDSINWQEE 11
   |||||
Db 53 RDSANWQAE 62

RESULT 23
Q6N9D6_RHOPA
ID Q6N9D6_RHOPA PRELIMINARY; PRT; 421 AA.
AC Q6N9D6_RHOPA
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative cytochrome P-450 (EC 1.14.-.-).
GN OrderedLocusNames=RPA1613;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres Y.Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris."
RL Nat. Biotechnol. 22:55-61(2004).
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; BX572598; CAB27054.1; -; Genomic DNA.
DR GO; GO:0046872; F-metal ion binding; IEA.
DR GO; GO:0004497; F-monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002397; BP450.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; p450; 1.

```



```
DR GO: GO:0030870; C:Mre11 complex; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0004519; F:endonuclease activity; IEA.
DR GO: GO:0006281; P:DNA repair; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003439; ABC transp_like.
DR InterPro: IPR004584; Rad50.
DR InterPro: IPR007517; Rad50 Zn hook.
DR Pfam: PF04423; Rad50_zn_hook; 1.
DR TIGRFAMs: TIGR00606; Repeat; Transport.
KW Complete proteome; Repeat; Transport.
SQ SEQUENCE 1296 AA; 149399 MW; 74987F28CDF37251 CRC64;

Query Match 49.0%; Score 47; DB 2; Length 1296;
Best Local Similarity 58.3%; Pred. No. 3e+02; Indels 0; Gaps 0;
Matches 7; Conservative 3; Mismatches 2;

QY 2 RDSIWPOEPPA 13
Db 157 QEDSLWPLSEPA 168
:|:|:|:|

RESULT 28
Q6KAF4 ORYSA PRELIMINARY; PRT; 159 AA.
ID O6KAF4_ ORYSA PRELIMINARY; PRT; 159 AA.
AC O6KAF4_
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Hypothetical protein OJ1046_F07.16.
GN Name=OJ1046_F07.16;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AF003984; BAD19110.1; -; Genomic_DNA.
DR Gramene; O6KAF4; -
KW Hypothetical protein.
SQ SEQUENCE 159 AA; 17734 MW; B7CCE57C844E7616 CRC64;

Query Match 47.9%; Score 46; DB 2; Length 159;
Best Local Similarity 43.8%; Pred. No. 43;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 RDSIWPOEPPAIRPR 17
Db 133 RGQOIWRRRRPSLEPR 148
:|:|:|:|

RESULT 29
Q7S4V0 NEUCR PRELIMINARY; PRT; 167 AA.
ID Q7S4V0_ NEUCR PRELIMINARY; PRT; 167 AA.
AC Q7S4V0_
DT 01-MAR-2004 (T-EMBLrel. 26, Created)
DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU02377.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehm B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thoman N., Barrett R., Gnerre S.,
RA Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Voelker R.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Radford A., Aramayo R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RA "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
RT Nature 0:0-0(2003).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
```

```
DR EMBL: AABX01000358; EAA30562.1; -; Genomic DNA.
SQ SEQUENCE 167 AA; 18497 MW; BCE9BE6302EE906 CRC64;

Query Match      47.9%; Score 46; DB 2; Length 167;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      6 IWQPEEPAIRPR 17
Db      :|||||
47 LWPQESSAIRTR 58

RESULT 30
Q6CV53_KLUJLA
ID Q6CV53_KLUJLA PRELIMINARY; PRT; 265 AA.
AC Q6CV53;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similarities with ca|CA3594|PF6076.3 Candida albicans unknown
DE function.
GN OrderedLocusNames=KLLA0A03113g;
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre B., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulou O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Sennen D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts";
RL Nature 430:35-44 (2004).
DR EMBL: CR382121; CAH02724.1; -; Genomic DNA.
KW Complete proteome.
SQ SEQUENCE 265 AA; 29436 MW; 5E1151BE00BA1289 CRC64;

Query Match      47.9%; Score 46; DB 2; Length 265;
Best Local Similarity 50.0%; Pred. No. 75;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1 FRDSSWPQEEPAIRP 16
Db      :|||||
18 FEDDSFVPQVTGPKL 33

RESULT 31
Q7QHW5_ANOGA
ID Q7QHW5_ANOGA PRELIMINARY; PRT; 830 AA.
AC Q7QHW5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP00000016022.
GN ORFNames=ENSANG00000013533;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAB01008811; EAA04913.2; -; Genomic DNA.
DR SNR; Q7QHW5; 130-654.
DR InterPro; IPR011989; ARM-like.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Arm; 6.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS00176; ARM_REPEAT; 7.
KW Cytoskeleton; Repeat.
SQ SEQUENCE 830 AA; 89308 MW; ED8F04D24C1E14D2 CRC64;

Query Match      47.9%; Score 46; DB 2; Length 830;
Best Local Similarity 56.2%; Pred. No. 2.6e+02;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      1 FRDSSWPQEEPAIRP 16
Db      :|||||
675 FRDSSWNNNGELGIGP 690

RESULT 32
WNK4_RAT
ID WNK4_RAT STANDARD; PRT; 1222 AA.
AC Q7TPR6; Q810H4; Q81LR5;
DT 23-MAR-2004 (Rel. 43, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Serine/threonine-protein kinase WNK4 [EC 2.7.1.37] (protein kinase
DE with no lysine 4) (protein kinase, lysine-deficient 4).
GN Names=Wnk4; Synonyms=PRKWNK4;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Wistar Kyoto; TISSUE=Kidney;
RX MEDLINE=22563432; PubMed=12642508;
RX DOI=10.1161/01.HYP.0000063147.92433.7D;
RA Monti J., Zimdahl H., Schulz H., Plehm R., Ganten D., Hubner N.;
RT "The role of Wnk4 in polygenic hypertension: a candidate gene analysis
RT on rat chromosome 10.";
RL Hypertension 41:938-942 (2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Wistar Kyoto; TISSUE=Kidney;
RX PubMed=14608358; DOI=10.1038/ng1271;
RA Kahle K.T., Wilson F.H., Leng Q., Lalloti M.D., O'Connell A.D.,
RA Dong K., Rapson A.K., MacGregor G.G., Giebisch G., Hebert S.C.,
RA Lifton R.P.;
RT "WNK4 regulates the balance between renal NaCl reabsorption and K+
RT secretion.";
RL Nat. Genet. 35:372-376 (2003).
CC -!- FUNCTION: Regulates the activity of the thiazide-sensitive Na-Cl
CC cotransporter, SLC12A3, by phosphorylation which appears to
CC prevent membrane trafficking of SLC12A3. Also inhibits the renal
CC K(+) channel, KCNJ1, via a kinase-independent mechanism by which
CC it induces clearance of the protein from the cell surface by
CC clathrin-dependent endocytosis. WNK4 appears to act as a molecular
CC switch that can vary the balance between NaCl reabsorption and
```

CC K(+) secretion to maintain integrated homeostasis (By similarity).  
CC -|- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -|- COFACTOR: Magnesium (By similarity).  
CC -|- ENZYME REGULATION: Activation requires autophosphorylation of Ser-493. Phosphorylation of Ser-489 also promotes increased activity (By similarity).  
CC -|- SUBUNIT: Interacts with the C-terminal region of KCNJ1 (By similarity).  
CC -|- SUBCELLULAR LOCATION: Present exclusively in intercellular junctions in the distal convoluted tubule and in both the cytoplasm and intercellular junctions in the cortical collecting duct. WNK4 is part of the tight junction complex (By similarity).  
CC -|- SIMILARITY: Belongs to the Ser/Thr protein kinase family. WNK subfamily.  
CC -|- CAUTION: Ref.1 sequence differs from that shown due to a frameshift in position 170.  
CC -|- CAUTION: Cys-361 is present instead of the conserved Lys which is expected to be an active site residue. Lys-344 appears to fulfill the required catalytic function.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
CC -----  
CC ENBL; AY187039; AAC18238.1; ALT FRAME; mRNA.  
CC ENBL; AY125677; AAC38858.1; -; mRNA.  
CC SMR; Q7TPK6; 322-591.  
CC Ensembl; ENSRNOG00000020441; Rattus norvegicus.  
CC RGD; 631401; Prkwnk4.  
CC GO; GO:0005923; C:tight junction; ISS.  
CC GO; GO:0005524; F:ATP binding; ISS.  
CC GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.  
CC GO; GO:0006811; P:ion transport; ISS.  
CC GO; GO:0006468; P:protein amino acid phosphorylation; ISS.  
CC GO; GO:0007243; P:protein kinase cascade; ISS.  
CC GO; GO:0050794; P:regulation of cellular process; ISS.  
CC InterPro; IPR008570; DUF852\_euk.  
CC InterPro; IPR000719; Prot\_kinase.  
CC InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
CC InterPro; IPR002290; Ser\_thr\_pkinase.  
CC Pfam; PF05871; DUF852\_1.  
CC Pfam; PF00069; Kinase\_1.  
CC ProDom; PD000001; Prot\_kinase; 1.  
CC SMART; SM00220; S\_TKc; 1.  
CC PROSITE; PS00107; PROTEIN KINASE ATP; FALSE\_NEG.  
CC PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.  
KW ATP-binding; Kinase; Nucleotide-binding; Phosphorylation;  
KW Serine/threonine-protein kinase; Transferase.  
FT DOMAIN 171 429 Protein kinase.  
FT NP\_BIND 177 185 ATP (By similarity).  
FT ACT\_SITE 299 299 Proton acceptor (By similarity).  
FT BINDING 183 183 ATP (By similarity).  
FT MOD\_RES 328 328 Phosphoserine (by autocatalysis) (By similarity).  
FT MOD\_RES 332 332 Phosphoserine (by autocatalysis) (By similarity).  
FT CONFLICT 9 9 T -> R (in Ref. 1).  
FT CONFLICT 204 204 T -> A (in Ref. 1).  
FT CONFLICT 764 764 P -> A (in Ref. 1).  
FT CONFLICT 776 776 A -> E (in Ref. 1).  
FT CONFLICT 1194 1194 R -> K (in Ref. 1).  
SQ SEQUENCE 1222 AA; 132833 MW; E7E53DAC9D1FA446 CRC64;  
Query Match 47.9%; Score 46; DB 1; Length 1222;  
Best Local Similarity 57.1%; Pred. No. 4e+02;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
Oy 3 DDSIWQPEEPAIRP 16  
Db 755 EDALSPQEEPAAMP 768  
:::|||||

## RESULT 33

Q6C910\_YARLI PRELIMINARY; PRT; 1292 AA.  
ID Q6C910;  
AC Q6C910;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Similar to sp|P12753 Saccharomyces cerevisiae YNL250w RAD50 DNA repair protein.  
DE protein.  
GN OrderedLocusNames=YALI0D15246g;  
OS Yarrowia lipolytica (Candida lipolytica).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Dipodascaceae; Yarrowia.  
OX NCBI\_TaxID=4954;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=CLIB 122 / E 150;  
RX PubMed=15229592; DOI=10.1038/nature02579;  
RA Dujon B., Sherman D., Fischer G., Durrans P., Casaregola S., Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talla E., Goffard N., Frangeul L., Aigie M., Anthouard V., Babour A., Barbe V., Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C., Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A., Despons L., Fabre B., Fairhead C., Ferry-Dumazet H., Groppi A., Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R., Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H., Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O., Pellenc S., Potier S., Richard G.-F., Straub M.-L., Suleau A., Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B., Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A., Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J., Wincker P., Souciet J.-L.;  
RT "Genome evolution in yeasts."  
RL Nature 430:35-44 (2004).  
DR ENBL; CR382130; CAG81040.1; -; Genomic\_DNA.  
DR GO; GO:0030870; C:Mre11 complex; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0016887; F:ATPase activity; IEA.  
DR GO; GO:0006281; P:DNA repair; IEA.  
DR GO; GO:0006810; P:transpos; IEA.  
DR InterPro; IPR003439; ABC transp\_like.  
DR InterPro; IPR004584; Rad50.  
DR TIGRFAMS; TIGR00606; rad50; 1.  
KW Complete proteome; Repeat; Transport.  
SQ SEQUENCE 1292 AA; 146628 MW; 43EE38D435DFEB7 CRC64;  
Query Match 47.9%; Score 46; DB 2; Length 1292;  
Best Local Similarity 46.2%; Pred. No. 4.2e+02;  
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;  
Oy 2 RDSIWQPEEPAI 14  
Db 157 QEESLWPLSEPAV 169  
:::|||||

RESULT 34  
Q9SB08\_ORYLO PRELIMINARY; PRT; 89 AA.  
ID Q9SB08\_ORYLO  
AC Q9SB08;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Catalase (EC 1.11.1.6) (Fragment).  
DE Name=cata2;  
GN Name=cata2;  
OS Oryza longistaminata (long-staminate rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4528;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.

```
RC STRAIN=C104977;
RA Iwamoto M., Maekawa M., Saito A., Higo H., Higo K.;
RT "Evolutionary relationship of plant catalase genes inferred from exon-
RT intron structures: isozyme divergence after the separation of monocots
RT and dicots.";
RL Theor. Appl. Genet. 97:9-19(1998).
DR EMBL; AB004771; BAA34211.1; -; Genomic_DNA.
DR HSSP; P46206; 1W7S.
DR Gramene; Q9S808; -.
DR GO; GO:0004096; F:catalase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0042744; P:hydrogen peroxide catabolism; IEA.
DR GO; GO:0006979; P:response to oxidative stress; IEA.
DR InterPro; IPR002226; Catalase.
DR InterPro; IPR011614; Catalase_N.
DR Pfam; PF00199; Catalase; 1.
DR ProDom; PD000510; Catalase; 1.
KW Oxidoreductase; Peroxidase.
FT NON_TER 1
FT NON_TER 89
SQ SEQUENCE 89 AA; 10180 MW; 3CE89CAACEFB0049 CRC64;

Query Match 46.9%; Score 45; DB 2; Length 89;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 DDSIWPOEPAIRP 16
DB 74 DTKTWPEEVLPRP 87

RESULT 35
Q9S7V3 ORYLO PRELIMINARY; PRT; 90 AA.
AC Q9S7V3
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 25, Last annotation update)
DE Catalase (EC 1.11.1.6) (Fragment).
GN Name-cata;
OS Oryza longistaminata (Long-staminate rice).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4528;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C105204; and C104075;
RA Iwamoto M., Maekawa M., Saito A., Higo H., Higo K.;
RT "Evolutionary relationship of plant catalase genes inferred from exon-
RT intron structures: isozyme divergence after the separation of monocots
RT and dicots.";
RL Theor. Appl. Genet. 97:9-19(1998).
DR EMBL; AB004775; BAA34215.1; -; Genomic_DNA.
DR EMBL; AB004773; BAA34213.1; -; Genomic_DNA.
DR HSSP; P46206; 1W7S.
DR Gramene; Q9S7V3; -.
DR GO; GO:0004096; F:catalase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0042744; P:hydrogen peroxide catabolism; IEA.
DR GO; GO:0006979; P:response to oxidative stress; IEA.
DR InterPro; IPR002226; Catalase.
DR InterPro; IPR011614; Catalase_N.
DR Pfam; PF00199; Catalase; 1.
DR ProDom; PD000510; Catalase; 1.
KW Oxidoreductase; Peroxidase.
FT NON_TER 1
FT NON_TER 90
SQ SEQUENCE 90 AA; 10336 MW; 3CBFEFDEB89573B CRC64;

Query Match 46.9%; Score 45; DB 2; Length 90;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 DDSIWPOEPAIRP 16
DB 74 DTKTWPEEVLPRP 87

RESULT 36
Q9S809 ORYLO PRELIMINARY; PRT; 90 AA.
AC Q9S809
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Catalase (EC 1.11.1.6) (Fragment).
GN Name-cata1;
OS Oryza longistaminata (Long-staminate rice).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4528;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C104977;
RA Iwamoto M., Maekawa M., Saito A., Higo H., Higo K.;
RT "Evolutionary relationship of plant catalase genes inferred from exon-
RT intron structures: isozyme divergence after the separation of monocots
RT and dicots.";
RL Theor. Appl. Genet. 97:9-19(1998).
DR EMBL; AB004770; BAA34210.1; -; Genomic_DNA.
DR HSSP; P46206; 1W7S.
DR Gramene; Q9S809; -.
DR GO; GO:0004096; F:catalase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0042744; P:hydrogen peroxide catabolism; IEA.
DR GO; GO:0006979; P:response to oxidative stress; IEA.
DR InterPro; IPR002226; Catalase.
DR InterPro; IPR011614; Catalase_N.
DR Pfam; PF00199; Catalase; 1.
DR ProDom; PD000510; Catalase; 1.
KW Oxidoreductase; Peroxidase.
FT NON_TER 1
FT NON_TER 90
SQ SEQUENCE 90 AA; 10336 MW; 146CE89CAACEFB00 CRC64;

Query Match 46.9%; Score 45; DB 2; Length 90;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 DDSIWPOEPAIRP 16
DB 74 DTKTWPEEVLPRP 87

RESULT 37
Q9ZWT5 ORYLO PRELIMINARY; PRT; 91 AA.
AC Q9ZWT5
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Catalase (EC 1.11.1.6) (Fragment).
GN Name-cata;
OS Oryza longistaminata (Long-staminate rice).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4528;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C105198;
```



```
RA Iwamoto M., Maekawa M., Saito A., Higo H., Higo K.;  
RT "Evolutionary relationship of plant catalase genes inferred from exon-  
intron structures: isozyme divergence after the separation of monocots  
and dicots";  
RL Theor. Appl. Genet. 97:9-19(1998).  
DR EMBL; AB004774; BAA34214.1; -; Genomic_DNA.  
DR HSSP; P46206; 1M7S.  
DR Gramene; Q9ZWT5; -.  
DR GO; GO:0004096; F:catalase activity; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR GO; GO:0042744; P:hydrogen peroxide catabolism; IEA.  
DR GO; GO:0006979; P:response to oxidative stress; IEA.  
DR InterPro; IPR002226; Catalase.  
DR InterPro; IPR011614; Catalase_N.  
DR Pfam; PF00199; Catalase; 1.  
DR ProDom; PD000510; Catalase; 1.  
KW Oxidoreductase; Peroxidase.  
FT NON_TER 1  
FT NON_TER 91  
SQ SEQUENCE 91 AA; 10493 MW; 560CBFEFEDEB8957 CRC64;  
  
Query Match 46.9%; Score 45; DB 2; Length 91;  
Best Local Similarity 50.0%; Pred. No. 34;  
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
  
QY 3 DDSIWPOEPAIRP 16  
Db 75 DTKTWPEEVPPLRP 88  
| :|:|:|:|  
RESULT 38  
Q5T558 HUMAN PRELIMINARY; PRT; 150 AA.  
AC Q5T558;  
DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
DE Katanin p60 subunit A-like 1 (Fragment).  
GN Name=KATNAL1; ORFNames=RP11-374F3.1-003;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Lovell J.;  
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL356750; CAL13720.1; -; Genomic_DNA.  
FT NON_TER 150  
SQ SEQUENCE 150 AA; 17091 MW; 4A66B556CC723944 CRC64;  
  
Query Match 46.9%; Score 45; DB 2; Length 150;  
Best Local Similarity 53.8%; Pred. No. 58;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 FRDPSIWQPEEPA 13  
Db 93 FRDPAVWPPVPVA 105  
| :|:|:|:|  
RESULT 39  
Q5JM49 ORYSA PRELIMINARY; PRT; 197 AA.  
AC Q5JM49;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DE Hypothetical protein P0042A10.36.  
GN Name=P0042A10.36;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI_TaxID=39947;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
RA Hishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
RA Ikono M., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,  
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
RA Yano M., Jiang J., Gojobori T.;  
RT "The genome sequence and structure of rice chromosome 1.";  
RL Nature 420:312-316(2002).  
DR EMBL; AP003343; BAD87458.1; -; Genomic_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 197 AA; 21575 MW; 8B6D745E2B699719 CRC64;  
  
Query Match 46.9%; Score 45; DB 2; Length 197;  
Best Local Similarity 42.9%; Pred. No. 78;  
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
QY 3 DDSIWPOEPAIRP 16  
Db 50 DDHLWPDPPRALP 63  
| :|:|:|:|  
RESULT 40  
Q7M7M9 WOLSU PRELIMINARY; PRT; 218 AA.  
AC Q7M7M9;  
DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE TRANSCRIPTIONAL REGULATOR (PHOB-LIKE).  
GN OrderedLocusNames=WS2201;  
OS Wolinella succinogenes.  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Helicobacteraceae; Wolinella.  
OX NCBI_TaxID=844;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=DSMZ 1740;  
RX MEDLINE=22882897; PubMed=14500908; DOI=10.1073/pnas.1932838100;  
RA Baar C., Eppinger M., Raddatz G., Simon J., Lanz C., Klimmek O.,  
RA Nandakumar R., Gross R., Rosinus A., Keller H., Jagtap P., Linke B.,  
RA Meyer F., Lederer H., Schuster S.C.;  
RT "Complete genome sequence and analysis of Wolinella succinogenes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695(2003).  
DR EMBL; BX571662; CAE11191.1; -; Genomic_DNA.  
DR HSSP; Q9WYN0; IKG5.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0000156; F:two-component response regulator activity; IEA.  
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.  
DR InterPro; IPR001789; Response_reg.  
DR InterPro; IPR001867; Trans_reg_C.  
DR InterPro; IPR011991; Wing_Hlx_DNA_bd.  
DR Pfam; PF000072; Response_reg; 1.  
DR Pfam; PF00486; Trans_reg_C; 1.  
DR ProDom; PD000039; Response_reg; 1.  
DR ProDom; PD000329; Trans_reg_C; 1.  
DR PROSITE; PSS0110; RESPONSE_REGULATORY; 1.  
KW Complete proteome.  
SQ SEQUENCE 218 AA; 24871 MW; 5C58F2C41E8229B6 CRC64;  
  
Query Match 46.9%; Score 45; DB 2; Length 218;
```



Best Local Similarity 60.0%; Pred. No. 87;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 DSIWPOEPA 13  
DB 177 DHLWPEEPS 186

## RESULT 41

Q5WS17 LEGPA  
ID Q5WS17 LEGPA PRELIMINARY; PRT; 415 AA.  
AC Q5WS17;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DE Hypothetical protein.  
GN ORFNAMES=plpp0130;  
OS Legionella pneumophila (strain Paris).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;  
OC Legionellaceae; Legionella.  
OX NCBI\_TaxID=297246;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Paris;  
RA Cazalet C., Rusniok C., Bruggemann H., Zidane N., Magnier A., Ma L.,  
RA Tichit M., Jarraud S., Bouchier C., Vandenesch F., Kunst F.,  
RA Etienne J., Glaeser P., Buchrieser C.;  
RT Evidence in the Legionella pneumophila genome for exploitation of  
RT host cell functions and high genome plasticity.";  
RL Nat. Genet. 36:1165-1173(2004).  
DR EMBL; CR28338; CAH17307.1; -; Genomic DNA.  
KW Hypothetical protein; Nucleotide-binding; Plasmid.  
SQ SEQUENCE 415 AA; 47681 MW; 6DD85749F4PFBBA CRC64;

Query Match 46.9%; Score 45; DB 2; Length 415;  
Best Local Similarity 50.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 DSIWPOEPAIRPR 17  
DB 351 DSLWNEEDNWVRPR 364

## RESULT 42

Q89E44 BRAJA  
ID Q89E44 BRAJA PRELIMINARY; PRT; 449 AA.  
AC Q89E44;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Blr7243 protein.  
GN OrderedLocustNames=blr7243;  
OS Bradyrhizobium japonicum.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Bradyrhizobium.  
OX NCBI\_TaxID=375;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=USDA 110;  
RX MEDLINE=22484998; PubMed=12597275;  
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
RA Sasanoto S., Watanabe A., Ideawa K., Iriiguchi M., Kawahima K.,  
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,  
RA Tabata S.;  
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
RT Bradyrhizobium japonicum USDA110.";  
RL DNA Res. 9:189-197(2002).

-!- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
monooxygenases. They oxidize a variety of structurally unrelated  
compounds, including steroids, fatty acids, and xenobiotics (By  
similarity).  
-!- SIMILARITY: Belongs to the cytochrome P450 family.

DR EMBL; BA000040; BAC52508.1; -; Genomic DNA.  
DR HSP; P3306; 1CPT.  
DR GO:0046872; F:metal ion binding; IEA.  
DR GO:0004497; F:monooxygenase activity; IEA.  
DR GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR002397; BP450.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PR00359; BP450.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN 1.  
KW Complete proteome; Heme; Iron; Metal-binding; Monooxygenase;  
KW Oxidoreductase.  
SQ SEQUENCE 449 AA; 51846 MW; 6FB122A36F2CCCOF CRC64;

Query Match 46.9%; Score 45; DB 2; Length 449;  
Best Local Similarity 44.4%; Pred. No. 1.9e+02;  
Matches 8; Conservative 4; Mismatches 0; Indels 6; Gaps 1;

QY 1 FRDSIWP-----QEEP 12  
DB 63 FQDDTLWPWFRLRKEEP 80

## RESULT 43

KATLI MOUSE  
ID KATLI MOUSE STANDARD; PRT; 488 AA.  
AC Q8K074;  
DT 10-MAY-2005 (Rel. 47, Created)  
DT 10-MAY-2005 (Rel. 47, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Katanin p60 ATPase-containing subunit A-like 1 (EC 3.6.4.3) (Katanin  
DE p60 subunit A-like 1) (p60 katanin-like 1).  
GN Name=Katanin;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muroidae; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=Eye;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

-!- FUNCTION: Severs microtubules in vitro in an ATP-dependent manner.  
This activity may promote rapid reorganization of cellular  
microtubule arrays (By similarity).  
-!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.  
-!- SIMILARITY: Belongs to the AAA ATPase family.

-!- This Swiss-Prot entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use as long as its content is in no way modified and this statement is not

```
CC removed.
CC -----
DR ENBL; BC030434; AAH30434.1; -; mRNA.
DR HSSP; Q01853; IR7R.
DR Ensembl; ENSMUSG000000041298; Mus musculus.
DR MGI; MGI:2387638; Katnall.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR003960; AAA_sub.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00674; AAA; 1.
KW ATP-binding; Hydrolase; Microtubule; Nucleotide-binding.
FT NP BIND 246 253 ATP (Potential).
SQ SEQUENCE 488 AA; 55165 MW; A9BFF34DF2F44BAC CRC64;

Query Match 46.9%; Score 45; DB 1; Length 488;
Best Local Similarity 53.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 FRDSIWPQEEPA 13
Db 93 FRDPAVWPFPVPA 105

RESULT 44
KATL1 RAT
ID KATL1 RAT STANDARD; PRT; 488 AA.
AC Q5XIK7;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Katanin p60 ATPase-containing subunit A-like 1 (EC 3.6.4.3) (Katanin
DE p60 subunit A-like 1) (p60 katanin-like 1).
GN Name=Katnall;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSU=Testis;
RG NIH - Mammalian Gene Collection (MGC) project;
RL Submitted (OCT-2004) to the ENBL/GenBank/DBJ databases.
CC -!- FUNCTION: Severs microtubules in vitro in an ATP-dependent manner.
CC This activity may promote rapid reorganization of cellular
CC microtubule arrays (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
CC -!- SIMILARITY: Belongs to the AAA ATPase family.
CC [2]
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR ENBL; BC030434; AAH30434.1; -; mRNA.
DR Ensembl; ENSRNOG0000000916; Rattus norvegicus.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR003960; AAA_sub.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00674; AAA; 1.
KW ATP-binding; Hydrolase; Microtubule; Nucleotide-binding.
FT NP BIND 246 253 ATP (Potential).
SQ SEQUENCE 488 AA; 55201 MW; 0ED5866744EEBCD6 CRC64;

Query Match 46.9%; Score 45; DB 1; Length 488;
Best Local Similarity 53.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 FRDSIWPQEEPA 13
Db 93 FRDPAVWPFPVPA 105

RESULT 45
KATL1 HUMAN
ID KATL1 HUMAN STANDARD; PRT; 490 AA.
AC Q9BM62;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Katanin p60 ATPase-containing subunit A-like 1 (EC 3.6.4.3) (Katanin
DE p60 subunit A-like 1) (p60 katanin-like 1).
GN Name=KATNAL1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OC NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC PubMed=15057823; DOI=10.1038/nature02379;
RA Dunham A., Matthews L.H., Burton J., Ashurst J.L., Howe K.L.,
RA Ashcroft K.J., Beare D.M., Burford D.C., Hunt S.E.,
RA Griffiths-Jones S., Jones M.C., Keenan S.J., Oliver K., Scott C.E.,
RA Ainscough R., Almeida J.P., Ambrose K.D., Andrews D.T.,
RA Ashwell R.I.S., Babbage A.K., Baggeley C.L., Bailey J., Bannerjee R.,
RA Barlow K.F., Bates K., Beasley H., Bird C.P., Bray-Allen S.,
RA Brown A.J., Brown J.Y., Burrill W., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M.E., Clark S.Y., Clarke G., Clee C.M.,
RA Clegg S.C., Cobley V., Collins J.E., Corby N., Coville G.J.,
RA Deloukas P., Dhami P., Dunham I., Dunn M., Earthrowl M.E.,
RA Ellington A.G., Faulkner L., Frankish A.G., Frankland J., French L.,
RA Garner P., Garnett J., Gilbert J.G.R., Gilson C.J., Ghorji J.,
RA Grafham D.V., Gribble S.M., Griffiths C., Hall R.E., Hammond S.,
RA Harley J.L., Hart E.A., Heath P.D., Howden P.J., Huckle E.J.,
RA Hunt P.J., Hunt A.R., Johnson C., Johnson D., Kay M., Kimberley A.M.,
RA King A., Laird G.K., Langford C.J., Lawlor S., Leongamornlert D.A.,
RA Lloyd D.M., Lloyd C., Loveland J.E., Lovell J., Martin S.,
RA Mashreghi-Mohammadi M., McLaren S.J., McMurray A., Milne S.,
RA Moore M.J.F., Nickerson T., Palmer S.A., Pearce A.V., Peck A.I.,
RA Pelan S., Phillimore B., Porter K.M., Rice C.M., Searle S.,
RA Sehra H.K., Showkeen R., Skuce C.D., Smith M., Steward C.A.,
RA Sycamore N., Tester J., Thomas D.W., Tracey A., Tromans A., Tubby B.,
RA Wall M., Wallis J.M., West A.P., Whitehead S.L., Willey D.L.,
RA Wilming L., Wray P.W., Wright M.W., Young L., Coulson A., Durbin R.,
RA Hubbard T., Sulston J.E., Beck S., Bentley D.R., Rogers J., Ross M.T.;
RA "The DNA sequence and analysis of human chromosome 13.";
RA Nature 428:522-528(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSU=Kidney;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
```

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Severs microtubules in vitro in an ATP-dependent manner.
CC This activity may promote rapid reorganization of cellular
CC microtubule arrays (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
CC -1- SIMILARITY: Belongs to the AAA ATPase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AL356750; CA113718.1; -; Genomic_DNA.
DR EMBL; BC000612; AAH00612.1; -; mRNA.
DR HSSP; Q01853; IR7R.
DR Ensembl; ENSG00000102781; Homo sapiens.
DR HGNC; HGNC:28361; KATNAL1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003960; AAA_sub.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00674; AAA; 1.
KW ATP-binding; Hydrolase; Microtubule; Nucleotide-binding.
FT NP_BIND 248 255 ATP (Potential).
SQ SEQUENCE 490 AA; 55392 MW; CB8E51A7F5CB487 CRC64;

Query Match 46.9%; Score 45; DB 1; Length 490;
Best Local Similarity 53.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 FRDSDSWPQEEPA 13
DB 93 FRDPAVWPBPVPA 105

RESULT 46
PKH1_YEAST
ID PKH1_YEAST STANDARD; PRT; 766 AA.
AC Q03407;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Serine/threonine-protein kinase PKH1 (BC 2.7.1.37) (3-phosphoinositide
DE dependent protein kinase-1).
GN Name=PKH1; OrderedLocNames=YDR490C;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S288c;
RX MEDLINE=97313263; PubMed=9169867;
RA Jacq C., Alt-Moerbe J., Andre B., Arnold W., Bahr A., Ballesta J.P.G.,
RA Barges M., Baron L., Becker A., Biteau N., Bloeker H., Blugnon C.,
RA Boskovic J., Brandt P., Brueckner M., Buitrago M.J., Coster F.,
RA Delaveau T., del Rey F., Dujon B., Eide L.G., Garcia-Cantalejo J.M.,
RA Goffeau A., Gomez-Peris A., Granotier C., Hanemann V., Hanelin T.,
RA Hoheisel J.D., Jaeger W., Jimenez A., Joniaux J.-L., Kraemer T.,
RA Kuester H., Laamanen P., Legros Y., Louis E.J., Moeller-Rieker S.,
RA Monnet A., Moro M., Mueller-Auer S., Nussbaumer B., Paricio N.,
RA Paulin L., Pera J., Perez-Alonso M., Perez-Ortin J.E., Pohl T.M.,
RA Prydz H., Purnelle B., Rasmussen S.W., Remacha M.A., Revuelta J.L.,
RA Rieger M., Salom D., Saluz H.P., Saiz J.E., Saren A.-M., Schaefer M.,
RA Scharte M., Schmidt E.R., Schneider C., Scholler P., Schwarz S.,
RA Soler-Mira A., Urrutazu L.A., Verhasselt P., Vissers S., Voet M.,
RA Volckaert G., Wagner G., Wambutt R., Wedler E., Wedler H., Woelfl S.,
RA Harris D.E., Bowman S., Brown D., Churche C.M., Connor R., Dedman K.,
RA Gentles S., Hamlin N., Hunt S., Jones L., McDonald S., Murphy L.D.,
RA Niblett D., Odell C., Oliver K., Rajandream M.A., Richards C.,
RA Shore L., Walsh S.V., Barrell B.G., Dietrich F.S., Mulligan J.T.,

```

```

RA Allen E., Araujo R., Aviles E., Berno A., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Hunnicke-Smith S., Hyman R.W.,
RA Komp C., Lashkari D., Lew H., Lin D., Mosedale D., Nakahara K.,
RA Namath A., Oefner P., Oh C., Petel F.X., Roberts D., Schramm S.,
RA Schroeder M., Shogren T., Shroff N., Winant A., Yelton M.A.,
RA Botstein D., Davis R.W., Johnston M., Andrews S., Brinkman R.,
RA Cooper J., Ding H., Du Z., Favello A., Fulton L., Gattung S.,
RA Greco T., Hallsworth K., Hawkins J., Hillier L.W., Jier M.,
RA Johnson D., Johnston L., Kirsten J., Kucaba T., Langston Y.,
RA Latreille P., Le T., Mardis E., Menezes S., Miller N., Nhan M.,
RA Pauley A., Peluso D., Rifkin L., Riles L., Taich A., Trevaaskis E.,
RA Vignati D., Wilcox L., Wohldman P., Vaubin M., Wilson R.,
RA Waterston R., Albermann K., Hani J., Heumann K., Kleine K.,
RA Mewes H.-W., Zollner A., Zaccaria P.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IV.";
RL Nature 387:75-78(1997).
RN [2]
RP FUNCTION
RX PubMed=10074427; DOI=10.1016/S0960-9822(99)80088-8;
RA Casamayor A., Torrance P.D., Kobayashi T., Thorner J., Alessi D.R.;
RT "Functional counterparts of mammalian protein kinases PDK1 and SGK in:
RT budding yeast.";
RL Curr. Biol. 9:186-197(1999).
CC -1- FUNCTION: Activates YPK1 by phosphorylating of a threonine
CC residue.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- INTERACTION:
CC P53252:YGR086C; NDEXP=1; IntAct=EBI-32467, EBI-23225;
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. PDK1
CC subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; U33050; AAB64917.1; -; Genomic_DNA.
DR PIR; S69657; S69657.
DR HSSP; O15530; 1H1W.
DR IntAct; Q03407; -
DR Ensembl; YDR490C; Saccharomyces cerevisiae.
DR SGD; S00002898; PKH1.
DR GO; GO:0005829; C:cytosol; IDA.
DR GO; GO:0004672; F:protein kinase activity; IDA.
DR GO; GO:0006897; P:endocytosis; IMP.
DR GO; GO:0000196; P:MAPKKK cascade during cell wall biogenesis; IGI.
DR GO; GO:0004668; P:protein amino acid phosphorylation; IDA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Complete proteome; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
FT DOMAIN 125 391 Protein kinase.
FT NP_BIND 131 139 ATP (By similarity).
FT ACT_SITE 249 249 Proton acceptor (By similarity).
FT BINDING 154 154 ATP (By similarity).
SQ SEQUENCE 766 AA; 86253 MW; 7FPD32B8BBBC367 CRC64;

Query Match 46.9%; Score 45; DB 1; Length 766;
Best Local Similarity 43.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 FRDSDSWPQEEPAIRP 16
DB 396 FEDGSVMDNDPPEIQP 411

```

```
RESULT 47
GSHAB_CLOPE
ID GSHAB_CLOPE STANDARD; PRT; 778 AA.
AC Q8XK30;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Glutathione biosynthesis bifunctional protein gshAB (Gamma-GCS-GS)
DE (GCS-GS) [includes: Glutamate--cysteine ligase (EC 6.3.2.2) (Gamma-
DE glutamylcysteine synthetase) (Gamma-ECS) (GCS); Glutathione synthetase
DE (EC 6.3.2.3) (Glutathione synthase) (GSH synthetase) (GSH-S) (GSHase)
DE (GS)].
GN Name=gshAB; OrderedLocusNames=CP151573;
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842; DOI=10.1073/pnas.022493799;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -!- FUNCTION: Synthesizes glutathione from L-glutamate and L-cysteine
CC via gamma-L-glutamyl-L-cysteine (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + L-cysteine = ADP +
CC phosphate + gamma-L-glutamyl-L-cysteine.
CC -!- CATALYTIC ACTIVITY: ATP + gamma-L-glutamyl-L-cysteine + glycine =
CC ADP + phosphate + glutathione.
CC -!- PATHWAY: Glutathione biosynthesis; first step.
CC -!- PATHWAY: Glutathione biosynthesis; second step.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SIMILARITY: In the N-terminal section; belongs to the glutamate--
CC cysteine ligase type 1 family, Type 2 subfamily.
CC -!- SIMILARITY: Contains 1 ATP-grasp domain.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BA000016; BAB1279.1; -; Genomic_DNA.
CC HAMAP; MF_00782; -; 1.
CC InterPro; IPR011761; ATP GRASP.
CC InterPro; IPR007370; Glu_cys_ligase.
CC InterPro; IPR006335; Glut_cys_rel.
CC InterPro; IPR002017; Spectrin.
CC DR PF04262; Glu_cys_ligase; 1.
CC DR TIGRFAMs; TIGR01435; glu_cys_lig_rel; 1.
CC DR PROSITE; PS50975; ATP GRASP; 1.
CC ATP-binding; Complete proteome; Glutathione biosynthesis; Ligase;
CC Multifunctional enzyme; Nucleotide-binding.
FT DOMAIN 521 777 ATP-grasp.
FT REGION 1 354 Glutamate--cysteine ligase.
SQ SEQUENCE 778 AA; 90463 MW; 2008FF46AF06423E CRC64;
Query Match 46.9%; Score 45; DB 1; Length 778;
Best Local Similarity 58.3%; Pred. No. 3.5e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 6 IWQEEPAIRP 17
Db 103 LWPQSNPILR 114
:|||||
:|||||
RESULT 48
Q56VR4_MOUSE
ID Q56VR4_MOUSE PRELIMINARY; PRT; 848 AA.
Query Match 46.9%; Score 45; DB 1; Length 848;
Best Local Similarity 57.1%; Pred. No. 3.8e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 3 DDSIWPOEEPAIRP 16
Db 755 EDALSPQEEPAALP 768
:|||||
:|||||
RESULT 49
Q4WCP6_ASPTU
ID Q4WCP6_ASPTU PRELIMINARY; PRT; 881 AA.
AC Q4WCP6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE GATA transcriptional activator (Area), putative.
GN ORFNames=Afu601970;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H., Stanley.,
RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Foster N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA Goldman G.H., Goni K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C.,
```

```
Q56VR4;
10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE WNK4 protein kinase.
DE Name=Prkwnk4;
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RA Mount D.B.;
RT "An alternatively spliced isoform of the mouse WNK4 protein kinase.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
EMBL; AY260059; AAP83467.1; -; mRNA.
DR SMR; Q56VR4; 161-430.
DR Ensembl; ENSMUSG0000035112; Mus musculus.
DR MGI; MGI:1917097; Prkwnk4.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0005923; C:tight junction; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0030644; P:chloride ion homeostasis; TAS.
DR GO; GO:0006821; P:chloride transport; IDA.
DR GO; GO:0008104; P:protein localization; IDA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase.
SQ SEQUENCE 848 AA; 93135 MW; 018A4D846366E76D CRC64;
```

RA Majoros W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,  
RA Mouyna I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,  
RA Penava M.A., Perte M., Price C., Pritchard B.L., Quail M.A.,  
RA Rabinowitz E., Rawlins N., Rajandream M.-A., Reichard U.,  
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,  
RA Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,  
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,  
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,  
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,  
RA Machida M., Hall N., Barrell B., Denning D.W.,  
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus  
RL *Aspergillus fumigatus*";  
CC Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL: AAHF01000012; EAL85942.1; -; Genomic DNA.  
SQ SEQUENCE 881 AA; 94519 MW; 7E94AE079546AAE3 CRC64;

Query Match 46.9%; Score 45; DB 2; Length 881;  
Best Local Similarity 66.7%; Pred. No. 4e+02;  
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 DDSIWPOEPAIRPR 17  
Db 185 DDFVIESPADRRP 199

RESULT 50  
Q6TAF9\_SOLTU  
ID Q6TAF9\_SOLTU PRELIMINARY; PRT; 948 AA.  
AC Q6TAF9;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Blight resistance protein SH10 (Fragment).  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
OC asterids; lamiids; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX PubMed=14675451; DOI=10.1046/j.1365-3113x.2003.01934.x;  
RA Van Der Vossen E., Sikkema A., Hekkert Bt B.L., Gros J., Stevens P.,  
RA Muskens M., Wouters D., Pereira A., Stiekema W., Allefs S.;  
RT "An ancient R gene from the wild potato species *Solanum bulbocastanum*  
RT confers broad-spectrum resistance to *Phytophthora infestans* in  
RT cultivated potato and tomato.";  
RL Plant J. 36:867-882(2003).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Van der Vossen E., Sikkema S.;  
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY426264; AAR29074.1; -; Genomic DNA.  
DR GO: GO:0005524; F:ATP binding; IEA.  
DR GO: GO:0006915; P:apoptosis; IEA.  
DR GO: GO:0042829; P:defense response to pathogen; IEA.  
DR InterPro: IPR000767; Disease resist.  
DR InterPro: IPR006209; EGF\_like.  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR002182; NB-ARC.  
DR Pfam: PF00560; LRR\_1; 7.  
DR Pfam: PF00931; NB-ARC; 1.  
DR PRINTS; PR00364; DISEASERSIST.  
DR PROSITE; PS01186; EGF\_2; UNKNOWN\_1.  
KW Plant defense.  
FT NON TER 948 948  
SQ SEQUENCE 948 AA; 108225 MW; 593E29C00EBCFC06 CRC64;

Query Match 46.9%; Score 45; DB 2; Length 948;  
Best Local Similarity 58.8%; Pred. No. 4.3e+02;  
Matches 10; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

Qy 2 RDDSIIW--POEPAIRP 16  
Db 374 RDSEIWNLPQERSILP 390

Search completed: February 3, 2006, 18:56:13  
Job time : 192.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:51:28 ; Search time 26 Seconds  
(without alignments)  
62.911 Million cell updates/sec

Title: US-10-693-538-1\_COPY\_46\_62

Perfect score: 96  
Sequence: 1 FRDSIWPOEPAIRPR 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 1000 summaries

Database :

PIR 80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	188	2 A39787	teratocarcinoma-de
2	96	100.0	188	2 A30362	teratocarcinoma-de
3	64	66.7	171	2 I49612	teratocarcinoma-de
4	48	50.0	496	2 S89027	protein T08H10.1 [
5	48	50.0	1081	2 S51899	probable protein k
6	45	46.9	766	2 S69657	hypothetical prote
7	44.5	46.4	906	2 T28034	hypothetical prote
8	44	45.8	352	2 T49396	hypothetical prote
9	44	45.8	1312	1 BWBYDL	RAD50 protein - ye
10	43	44.8	197	2 B95391	protein [imported
11	43	44.8	335	2 S70671	lipopolysaccharide
12	43	44.8	589	2 B90605	hypothetical prote
13	42	43.8	291	1 EEW7G	gamma-gliadin B pr
14	42	43.8	311	2 G87273	oxidoreductase ito
15	42	43.8	491	1 CSRZ	catalase (EC 1.1.1.
16	42	43.8	681	2 F64889	membrane protein m
17	42	43.8	831	2 S50163	nitrate reductase
18	42	43.8	1736	2 A17747	tight junction pro
19	41.5	43.2	503	2 S14775	steroid 6beta-mono
20	41	42.7	78	2 C84085	hypothetical prote
21	41	42.7	152	2 E87272	hypothetical prote
22	41	42.7	223	2 F72803	gp32 protein - Myc
23	41	42.7	279	1 SNPF5K	proteasome endopep
24	41	42.7	327	2 C83883	probable alpha-N-a
25	41	42.7	540	2 T50617	hypothetical prote
26	41	42.7	544	2 C95854	probable trehalose
27	41	42.7	606	2 H97012	hypothetical prote
28	41	42.7	639	2 H70661	probable dnaG prot
29	41	42.7	642	2 C87013	DNA primase [impor

30	41	42.7	658	2 T50080	probable DNA repai
31	41	42.7	721	2 S49789	hypothetical prote
32	41	42.7	756	2 T34797	probable ATP bindi
33	41	42.7	828	2 G91015	probable nitrate r
34	41	42.7	828	2 A85860	probable nitrate r
35	41	42.7	1214	2 T47438	disease resistance
36	40.5	42.2	124	2 A86612	cytochrome P450 3A
37	40.5	42.2	488	2 A82115	sigma-54 dependent
38	40	41.7	240	2 AG0605	oxygen-insensitive
39	40	41.7	240	2 I80318	drug activity modu
40	40	41.7	240	2 C90745	modulator of drug
41	40	41.7	240	2 G85595	modulator of drug
42	40	41.7	249	2 E72703	probable reductase
43	40	41.7	312	1 S74432	hypothetical prote
44	40	41.7	324	2 AG2956	hypothetical prote
45	40	41.7	324	2 G98326	c4-dicarboxylate-b
46	40	41.7	328	2 C97268	characterized pr
47	40	41.7	353	2 T01256	probable anthocyan
48	40	41.7	492	2 A55589	lysine-tRNA ligase
49	40	41.7	617	2 T36294	dihydroxy-acid deh
50	40	41.7	701	2 T16607	hypothetical prote
51	40	41.7	1298	2 T24480	hypothetical prote
52	40	41.7	1305	2 D82923	hypothetical prote
53	40	41.7	2262	2 T30890	DNA-directed RNA p
54	39	40.6	29	2 S70328	calcium channel al
55	39	40.6	108	1 K1HUWE	gamma35 secalin -
56	39	40.6	206	2 S40829	hypothetical 23.5K
57	39	40.6	206	2 G86076	probable phosphata
58	39	40.6	206	2 H91229	probable phosphata
59	39	40.6	225	2 H83750	hypothetical prote
60	39	40.6	227	2 AH3411	hypothetical prote
61	39	40.6	244	2 F96733	hypothetical prote
62	39	40.6	251	2 PS0094	gamma-gliadin prec
63	39	40.6	319	2 A40619	lipopolysaccharide
64	39	40.6	330	2 D86038	heptosyl transfera
65	39	40.6	330	2 C91191	heptosyl transfera
66	39	40.6	429	2 JC4965	eik1 protein - mou
67	39	40.6	429	2 F86240	hypothetical prote
68	39	40.6	476	2 T16686	hypothetical prote
69	39	40.6	521	2 T11166	CDPdiacylglycerol-
70	39	40.6	531	2 T44803	type I site-specif
71	39	40.6	531	2 T09460	type I site-specif
72	39	40.6	531	2 T30323	type I site-specif
73	39	40.6	553	1 A42499	mullerian inhibiti
74	39	40.6	575	2 T08964	hypothetical prote
75	39	40.6	667	2 G90883	probable collagena
76	39	40.6	667	2 H85734	probable collagena
77	39	40.6	667	2 F64895	hypothetical prote
78	39	40.6	673	2 T46340	hypothetical prote
79	39	40.6	692	2 T15819	hypothetical prote
80	39	40.6	849	1 S64732	scaffold attachmen
81	39	40.6	984	2 T44496	cellulose 1,4-beta
82	39	40.6	1028	2 A59253	myosin I beta - hu
83	39	40.6	1254	2 S46636	hypothetical prote
84	39	40.6	1734	2 A54602	microtubule-associ
85	38.5	40.1	233	2 AD3344	hypothetical prote
86	38.5	40.1	339	2 A97297	NADH:flavin oxidor
87	38.5	40.1	488	2 A87660	hypothetical prote
88	38.5	40.1	570	1 A36954	hydroxylamine oxid
89	38.5	40.1	1644	2 AC0823	probable lipoprote
90	38.5	40.1	1653	2 B91052	hypothetical prote
91	38.5	40.1	1653	2 F85896	hypothetical prote
92	38.5	40.1	1653	2 G65028	hypothetical prote
93	38	39.6	23	2 S70327	gamma70 secalin -
94	38	39.6	30	2 S52126	gamma3-gliadin P25
95	38	39.6	92	2 B81235	preprotein translo
96	38	39.6	92	2 B82008	probable preprotei
97	38	39.6	104	2 F70014	conserved hypotet
98	38	39.6	104	2 AE1430	B. subtilis YulD p
99	38	39.6	104	2 AC1804	B. subtilis YulD p
100	38	39.6	129	2 T26423	hypothetical prote
101	38	39.6	131	2 T01328	profilin 4 - maize
102	38	39.6	136	2 T48713	hypothetical prote

103	38	39.6	136	2	C72644	hypothetical prote	176	37	38.5	120	2	A87462	hypothetical prote
104	38	39.6	206	2	T29985	hypothetical prote	177	37	38.5	125	2	F87685	hypothetical prote
105	38	39.6	222	2	A75406	hydrolase - Deinoc	178	37	38.5	130	2	AH0463	probable acetyltra
106	38	39.6	293	2	T37227	probable sugar tra	179	37	38.5	147	2	JQ1134	hypothetical prote
107	38	39.6	296	2	S24386	protein kinase (EC	180	37	38.5	149	2	G87329	hypothetical prote
108	38	39.6	298	2	H83163	probable transcrip	181	37	38.5	177	2	T31086	probable drpp-4-de
109	38	39.6	302	2	JA0153	gamma-gliadin prec	182	37	38.5	199	2	AF0947	probable haloacid
110	38	39.6	308	1	CRH06	carbonate dehydrat	183	37	38.5	201	2	A82466	hypothetical prote
111	38	39.6	316	2	T14643	1-aminocyclopropan	184	37	38.5	206	2	B48441	antigen (C-termina
112	38	39.6	329	2	A29894	isopenicillin N sy	185	37	38.5	239	2	T39885	hypothetical prote
113	38	39.6	357	2	S33144	leucoanthocyanidin	186	37	38.5	244	2	T25611	hypothetical prote
114	38	39.6	373	2	T10577	hypothetical prote	187	37	38.5	245	2	S19018	complement subcomp
115	38	39.6	374	2	AF1280	probable exonuclea	188	37	38.5	255	2	A60637	merozoite antigen
116	38	39.6	374	2	AF1643	probable exonuclea	189	37	38.5	256	2	T36557	probable hydrolase
117	38	39.6	383	2	T06238	omega-3 fatty acid	190	37	38.5	258	2	S75950	hypothetical prote
118	38	39.6	432	2	E96712	unknown protein, 6	191	37	38.5	271	2	C82104	vibriobactin utili
119	38	39.6	438	2	D84331	hypothetical prote	192	37	38.5	273	2	AH2432	hypothetical prote
120	38	39.6	445	2	A75376	probable oligosacc	193	37	38.5	278	2	S73153	hypothetical prote
121	38	39.6	469	2	F72535	hypothetical prote	194	37	38.5	283	2	H64764	probable taurine d
122	38	39.6	478	2	JC4838	bone morphogenetic	195	37	38.5	283	2	F90681	taurine dioxygenas
123	38	39.6	496	2	S51668	tyrosine kinase -	196	37	38.5	283	2	B85532	taurine dioxygenas
124	38	39.6	499	2	T48755	probable PELI (pho	197	37	38.5	314	2	AB2409	dTDP-glucose dehyd
125	38	39.6	513	2	C86897	hypothetical prote	198	37	38.5	327	2	T50744	spheroidene monoox
126	38	39.6	529	2	A47082	probable transcrip	199	37	38.5	327	2	S49619	crtA protein - Rho
127	38	39.6	529	2	A12119	transcription regu	200	37	38.5	328	2	S15284	isopenicillin N sy
128	38	39.6	530	2	S46322	phosphoprotein pho	201	37	38.5	346	2	T44019	hypothetical prote
129	38	39.6	534	2	G90925	hypothetical prote	202	37	38.5	348	2	C96511	probable heat shoc
130	38	39.6	534	2	C85774	hypothetical prote	203	37	38.5	350	2	T44206	hypothetical prote
131	38	39.6	534	2	D64924	hypothetical prote	204	37	38.5	355	2	H96834	hypothetical prote
132	38	39.6	555	1	S20100	hypothetical prote	205	37	38.5	379	2	G70918	hypothetical prote
133	38	39.6	569	2	B99587	mullerian inhibiti	206	37	38.5	425	2	T24522	hypothetical prote
134	38	39.6	571	2	S52276	translation initia	207	37	38.5	436	2	B55452	cartilage-derived
135	38	39.6	577	2	B75585	probable long-chai	208	37	38.5	437	2	D83329	conserved hypotet
136	38	39.6	652	2	A27323	parasporal crystal	209	37	38.5	441	2	A43555	GAP-43-related pro
137	38	39.6	730	2	T22517	probable procollag	210	37	38.5	470	2	S54029	hypothetical prote
138	38	39.6	743	2	F71062	hypothetical prote	211	37	38.5	472	2	AC0863	L-fucose kinase
139	38	39.6	775	2	S67920	DNA-directed DNA p	212	37	38.5	485	2	B40552	bindin fertilizati
140	38	39.6	798	2	T00131	xylan 1,4-beta-xy	213	37	38.5	487	2	B83882	sucrase-6-phosphat
141	38	39.6	807	2	H75634	myosin-1c - mouse	214	37	38.5	513	2	S75604	panthothenate synth
142	38	39.6	824	2	AD3098	periplasmic nitrat	215	37	38.5	529	2	T08684	hypothetical prote
143	38	39.6	828	2	D64990	probable nitrate r	216	37	38.5	531	2	T02952	probable protein p
144	38	39.6	828	2	T06133	hypothetical prote	217	37	38.5	538	2	S22409	D-alanyl-D-alanine
145	38	39.6	834	2	E98188	periplasmic nitrat	218	37	38.5	570	2	T38148	phosphatidyl synth
146	38	39.6	836	2	AB0540	outer-membrane fim	219	37	38.5	578	2	T33916	hypothetical prote
147	38	39.6	909	2	C84965	oxoglutarate dehyd	220	37	38.5	593	2	T01575	beta-fructofuranos
148	38	39.6	976	2	B87291	2-oxoglutarate deh	221	37	38.5	594	2	I49127	intracellular prot
149	38	39.6	988	2	D97600	excinuclease ABC c	222	37	38.5	629	2	T37255	acetylcholinestera
150	38	39.6	1006	2	D86431	protein T518.6 [im	223	37	38.5	635	2	A81656	threonyl-tRNA synt
151	38	39.6	1028	2	S41749	myosin heavy chain	224	37	38.5	656	2	E75468	hypothetical prote
152	38	39.6	1028	2	S37146	myosin I heavy cha	225	37	38.5	660	2	G71497	threonine-tRNA lig
153	38	39.6	1040	2	T29092	TSC-22 protein hom	226	37	38.5	673	1	HJECDR	ATP-dependent DNA
154	38	39.6	1178	2	S76370	sensory transducti	227	37	38.5	673	2	H86063	hypothetical prote
155	38	39.6	1292	2	D84727	probable RAD50 DNA	228	37	38.5	673	2	G91217	ATP-dependent DNA
156	38	39.6	1442	2	T42607	transcription acti	229	37	38.5	674	2	AI0922	hypothetical prote
157	38	39.6	1487	1	EDBEP1	immediate-early pr	230	37	38.5	686	2	S66693	hypothetical prote
158	38	39.6	1487	1	EDBEP6	155K transcription	231	37	38.5	702	2	C84111	hypothetical prote
159	38	39.6	1964	2	T09059	notch4 - mouse	232	37	38.5	751	2	F83080	hypothetical prote
160	38	39.6	3343	2	T42207	breat cancer susc	233	37	38.5	776	2	A87330	sensory box histid
161	37.5	39.1	129	2	H72711	hypothetical prote	234	37	38.5	829	2	D82430	periplasmic nitrat
162	37.5	39.1	364	2	T30723	hypothetical prote	235	37	38.5	843	2	T46717	hypothetical prote
163	37.5	39.1	376	2	E84742	hypothetical prote	236	37	38.5	851	2	T38497	hypothetical prote
164	37.5	39.1	380	2	E84644	hypothetical prote	237	37	38.5	861	2	C87714	leucyl-tRNA synthe
165	37.5	39.1	406	2	AG2704	MFS permease limpo	238	37	38.5	907	2	T02417	probable C2H2-type
166	37.5	39.1	414	2	G97486	probable mfs trans	239	37	38.5	933	1	DEECOG	oxoglutarate dehyd
167	37.5	39.1	430	2	T46317	hypothetical prote	240	37	38.5	933	2	AD0591	2-oxoglutarate deh
168	37.5	39.1	488	2	S71029	flaK protein - Vib	241	37	38.5	933	2	F85573	hypothetical prote
169	37.5	39.1	523	2	AG0984	conserved hypotet	242	37	38.5	933	2	G90722	oxoglutarate dehyd
170	37.5	39.1	594	2	G83878	L-lactate permease	243	37	38.5	935	2	AI0136	oxoglutarate dehyd
171	37.5	39.1	1687	2	T30176	Egr repeat transme	244	37	38.5	936	2	B82121	2-oxoglutarate deh
172	37	38.5	53	2	F70956	hypothetical prote	245	37	38.5	950	2	B64135	hypothetical prote
173	37	38.5	87	2	S72830	hypothetical prote	246	37	38.5	965	2	T12545	oxoglutarate dehyd
174	37	38.5	108	1	K1HUGL	Ig kappa chain V-I	247	37	38.5	967	2	F82668	oxoglutarate dehyd
175	37	38.5	108	2	S36279	Ig lambda chain V	248	37	38.5	998	2	T04214	hypothetical prote



249	37	38.5	1057	2	T15720	hypothetical prote	322	36	37.5	420	2	A87327	sodium-dicarboxyla
250	37	38.5	1159	2	T43461	probable phosphodi	323	36	37.5	429	2	A60145	hemoglobinase (EC
251	37	38.5	1237	2	T45070	protein kinase hom	324	36	37.5	437	2	D69519	tRNA nucleotidyltr
252	37	38.5	1529	2	T20986	hypothetical prote	325	36	37.5	438	2	T16037	hypothetical prote
253	37	38.5	1687	2	T30244	phosphodiesterase	326	36	37.5	438	2	S25260	porin precursor, p
254	37	38.5	1706	2	T30175	exoribonuclease, v	327	36	37.5	449	1	NBHUS	complement factor
255	37	38.5	1719	2	T30174	exoribonuclease, v	328	36	37.5	449	2	S27490	catalase (EC 1.11.
256	37	38.5	1854	2	T39035	putative sodium ch	329	36	37.5	451	2	T36149	hypothetical prote
257	37	38.5	2073	2	T39207	fatty acid synthas	330	36	37.5	469	2	A99656	hypothetical prote
258	37	38.5	2073	2	T43311	fatty-acyl-CoA syn	331	36	37.5	473	2	JQ1936	xylan 1,4-beta-xy
259	37	38.5	2701	2	T17796	inositol-trisphosp	332	36	37.5	475	2	T74780	hypothetical prote
260	37	38.5	7576	2	T17428	PK506 polyketide s	333	36	37.5	496	2	F87423	ferredoxin, probab
261	36.5	38.0	261	2	S76484	hypothetical prote	334	36	37.5	496	2	AB2432	D-hydantoinase [im
262	36.5	38.0	361	2	T21336	hypothetical prote	335	36	37.5	499	2	B85507	hypothetical prote
263	36.5	38.0	474	1	YFBXAM	phenylalanine-tRNA	336	36	37.5	502	2	AH2289	4-alpha-glucanotra
264	36.5	38.0	615	2	T41848	threonyl-tRNA synt	337	36	37.5	511	2	AB0858	hypothetical prote
265	36.5	38.0	661	2	T48943	hypothetical prote	338	36	37.5	512	2	D82444	NAD(P) transhydrog
266	36.5	38.0	1736	2	T05174	hypothetical prote	339	36	37.5	535	2	F70782	probable oxidoredu
267	36.5	38.0	3512	2	T17121	CPY protein - midg	340	36	37.5	544	2	A75414	B-3-phosphoglycra
268	36	37.5	79	2	F72354	hypothetical prote	341	36	37.5	556	2	A44441	B-cell antigen CD1
269	36	37.5	81	2	B95286	hypothetical prote	342	36	37.5	571	2	T18116	hypothetical prote
270	36	37.5	86	2	S16836	Ig kappa chain V r	343	36	37.5	591	2	H87338	sensor histidine k
271	36	37.5	87	2	B37475	probable structura	344	36	37.5	593	2	S45281	coagulation factor
272	36	37.5	94	2	A83108	hypothetical prote	345	36	37.5	628	2	AF3438	heavy metal tolera
273	36	37.5	108	1	K1HUB1	Ig kappa chain V-1	346	36	37.5	631	2	AH0315	probable ABC trans
274	36	37.5	116	2	B25521	Ig kappa chain pre	347	36	37.5	638	1	YBSST2	threonine-tRNA lig
275	36	37.5	123	2	S40313	Ig kappa chain V-J	348	36	37.5	648	2	JCS930	serine/arginine-ri
276	36	37.5	145	2	T14411	photosystem I prot	349	36	37.5	649	1	JH0261	parapaporal crystal
277	36	37.5	149	2	T12930	hypothetical prote	350	36	37.5	671	2	A45720	phenol 2-monooxyge
278	36	37.5	153	2	E89794	conserved hypotet	351	36	37.5	672	2	G88651	protein B0212.3 [1
279	36	37.5	165	2	E91141	stringent starvati	352	36	37.5	719	2	JC1200	alpha-glucosidase
280	36	37.5	165	2	J05666	stringent starvati	353	36	37.5	767	2	B95301	conserved hypotet
281	36	37.5	165	2	H85986	stringent starvati	354	36	37.5	774	2	AB6454	hypothetical prote
282	36	37.5	165	2	C72116	hypothetical prote	355	36	37.5	776	2	AB1796	gamma-glutamylcyst
283	36	37.5	176	2	A86746	hypothetical prote	356	36	37.5	776	2	AI1420	gamma-glutamylcyst
284	36	37.5	196	2	S17703	glutamine amidotra	357	36	37.5	818	2	H83904	hypothetical prote
285	36	37.5	199	2	AB0798	probable oxetanoci	358	36	37.5	828	2	AI0788	probable nitrate r
286	36	37.5	212	2	S74376	hypothetical prote	359	36	37.5	829	2	F83499	periplasmic nitrat
287	36	37.5	214	2	S15326	fimbrial protein M	360	36	37.5	831	2	A48489	nitrate reductase
288	36	37.5	217	2	JQ1294	hypothetical 23.7K	361	36	37.5	893	2	T15183	hypothetical prote
289	36	37.5	224	2	T12332	glutathione transf	362	36	37.5	1006	2	T31685	beta-galactosidase
290	36	37.5	226	2	I39594	exEB protein - Aer	363	36	37.5	1023	2	S12519	glutactin - fruit
291	36	37.5	233	2	S28356	hypothetical prote	364	36	37.5	1029	1	S21369	multidrug efflux t
292	36	37.5	237	2	G72129	hypothetical prote	365	36	37.5	1029	2	C97665	acriflavin resista
293	36	37.5	237	2	H86491	hypothetical prote	366	36	37.5	1029	2	AF2889	hypothetical prote
294	36	37.5	238	2	AI2178	hypothetical prote	367	36	37.5	1048	2	F70592	myosin-If - mouse
295	36	37.5	248	2	T26523	hypothetical prote	368	36	37.5	1099	2	AS9300	hypothetical membr
296	36	37.5	251	2	T39332	hypothetical prote	369	36	37.5	1226	2	AB3327	complement factor
297	36	37.5	261	2	D98256	hypothetical prote	370	36	37.5	1231	1	NBHUH	retrovirus-related
298	36	37.5	261	2	AG3028	hypothetical prote	371	36	37.5	1268	2	S52781	hypothetical TPR d
299	36	37.5	267	2	B81541	hypothetical prote	372	36	37.5	1317	2	B41950	hypothetical prote
300	36	37.5	270	2	G96914	chemotaxis protein	373	36	37.5	1389	2	T41230	probable ATP-depen
301	36	37.5	275	2	F81675	conserved hypotet	374	36	37.5	1389	2	B90924	ATP-dependent helli
302	36	37.5	295	2	T04483	probable ring fing	375	36	37.5	1538	2	F85772	probable ATP-depen
303	36	37.5	296	2	S72601	hypothetical prote	376	36	37.5	2037	2	S37178	fatty-acyl-CoA syn
304	36	37.5	305	2	H75091	hypothetical prote	377	36	37.5	2403	2	AS9386	sanko - human
305	36	37.5	308	2	A69318	hypothetical prote	378	36	37.5	343	2	S44887	ZK112.7 protein -
306	36	37.5	311	2	S72942	tRNA isopentenyltr	379	36	37.5	343	2	S44887	cyclosporin synthe
307	36	37.5	311	2	A48441	antigen (C-termina	380	36	37.5	15281	2	S41309	hypothetical prote
308	36	37.5	322	2	S60783	lipopolysaccharide	381	35.5	37.0	261	2	T29113	probable inositol
309	36	37.5	324	2	H95894	probable periplasm	382	35.5	37.0	501	2	H84727	hypothetical prote
310	36	37.5	340	2	E72739	conserved hypotet	383	35.5	37.0	509	2	T04688	hypothetical prote
311	36	37.5	350	2	G33560	6R protein - human	384	35.5	37.0	514	2	D84723	hypothetical prote
312	36	37.5	355	2	T10722	anthocyanidin synt	385	35.5	37.0	553	2	T04683	hypothetical prote
313	36	37.5	359	2	S22505	peroxidase (EC 1.1	386	35.5	37.0	647	2	T43952	hypothetical prote
314	36	37.5	365	2	T08679	procollagen-lysine	387	35.5	37.0	759	2	T44142	hypothetical prote
315	36	37.5	367	2	T22663	hypocotaxin	388	35.5	37.0	1092	2	S42798	DR1 protein [impor
316	36	37.5	375	2	T03593	leucoanthocyanidin	389	35.5	37.0	1323	2	I78557	fibronectin-bindin
317	36	37.5	386	2	T15301	hypothetical prote	390	35.5	37.0	1356	1	C45219	N-methyl-D-asparta
318	36	37.5	394	2	B86507	hypothetical prote	391	35	36.5	25	2	S09540	gene 1.5 protein -
319	36	37.5	394	2	B81554	hypothetical prote	392	35	36.5	73	2	G84360	hypothetical prote
320	36	37.5	400	2	D82953	conserved hypotet	393	35	36.5	88	2	T29934	hypothetical prote
321	36	37.5	405	2	T07699	hypothetical prote	394	35	36.5	98	2	PH1074	Ig light chain V r

395	35	36.5	99	2	JH0429	Ig kappa chain V-I	468	35	36.5	472	2	S53866	cytosine methylase
396	35	36.5	104	2	E42074	hypothetical prote	469	35	36.5	474	2	T48781	MEIOTIC RECOMBINAT
397	35	36.5	109	2	T21958	hypothetical prote	470	35	36.5	476	2	T30781	kelch protein - homo
398	35	36.5	110	2	FN0535	Ig kappa chain V r	471	35	36.5	480	2	H42506	F3L protein - vacc
399	35	36.5	111	2	C72714	hypothetical prote	472	35	36.5	480	2	G36213	F3L protein - vacc
400	35	36.5	114	2	S46377	Ig kappa chain V-J	473	35	36.5	481	2	S39682	ywL protein - Bac
401	35	36.5	114	2	T23119	hypothetical prote	474	35	36.5	483	2	AE3480	aminocyclase (EC 3
402	35	36.5	116	2	PT0357	Ig kappa chain V r	475	35	36.5	484	2	S61870	hypothetical prote
403	35	36.5	118	2	T03036	Ig light chain - h	476	35	36.5	492	2	A31120	phytoene dehydroge
404	35	36.5	125	2	E81783	hypothetical prote	477	35	36.5	498	2	S01113	alpha-amylase (EC
405	35	36.5	125	2	H81206	hypothetical prote	478	35	36.5	510	2	T05314	hypothetical prote
406	35	36.5	126	2	S40335	Ig kappa chain V-J	479	35	36.5	511	2	T04221	hypothetical prote
407	35	36.5	127	2	S09230	Ig kappa chain pre	480	35	36.5	512	2	H86216	protein T23G18.16
408	35	36.5	128	2	A64335	conserved hypotet	481	35	36.5	536	2	B75329	catalase (EC 1.11.
409	35	36.5	128	2	E81206	hypothetical prote	482	35	36.5	536	2	T05523	cytochrome P450 mo
410	35	36.5	135	2	A81207	hypothetical prote	483	35	36.5	543	2	T18694	hypothetical prote
411	35	36.5	150	2	JH0249	guanine nucleotide	484	35	36.5	551	2	T06777	phosphoinositide-s
412	35	36.5	153	2	A81048	conserved hypotet	485	35	36.5	572	2	S46712	heat shock protein
413	35	36.5	153	2	H86112	hypothetical prote	486	35	36.5	592	1	A30314	protein kinase C (
414	35	36.5	153	2	S56393	hypothetical 16.9K	487	35	36.5	592	1	JC1480	protein kinase C (
415	35	36.5	153	2	H91271	hypothetical prote	488	35	36.5	592	2	JN0877	protein kinase C (
416	35	36.5	187	2	AG2619	conserved hypotet	489	35	36.5	594	2	T06467	phosphoinositide-s
417	35	36.5	187	2	G97401	hira-interacting p	490	35	36.5	594	2	S33561	ref(2)P protein -
418	35	36.5	190	2	A13299	5-methylcytosine-s	491	35	36.5	596	2	G87659	peptidase M3 famil
419	35	36.5	209	2	E82241	preprotein translo	492	35	36.5	597	2	C75351	translation initia
420	35	36.5	219	2	G82611	hypothetical prote	493	35	36.5	605	2	S63370	probable membrane
421	35	36.5	221	2	H86307	probable glutathio	494	35	36.5	609	2	T40660	hypothetical C2H2-
422	35	36.5	234	2	T25724	hypothetical prote	495	35	36.5	610	2	S19461	probable membrane
423	35	36.5	240	2	D82288	NADPH-flavin oxido	496	35	36.5	622	2	E71312	hypothetical prote
424	35	36.5	240	2	G83271	probable phosphodi	497	35	36.5	643	1	F0LJLK	gag polyprotein -
425	35	36.5	246	2	F72766	hypothetical prote	498	35	36.5	646	1	QOEC85	hypothetical 73.3K
426	35	36.5	248	1	JQ1682	infected cell prot	499	35	36.5	646	2	T26427	hypothetical prote
427	35	36.5	258	2	AF0016	probable biotin bi	500	35	36.5	646	2	D91144	hypothetical prote
428	35	36.5	269	2	G75558	hypothetical prote	501	35	36.5	654	2	A80989	hypothetical prote
429	35	36.5	270	2	B86486	protein F28J9.7 [i	502	35	36.5	654	2	GH0668	probable peptidase
430	35	36.5	272	2	T25874	hypothetical prote	503	35	36.5	660	2	E81839	probable transmemb
431	35	36.5	272	2	F71162	hypothetical prote	504	35	36.5	673	2	AG0471	ATP-dependent DNA
432	35	36.5	276	2	D72001	hypothetical prote	505	35	36.5	676	2	T34609	probable helicase
433	35	36.5	276	2	D86622	hypothetical prote	506	35	36.5	691	2	C84300	cell division cycl
434	35	36.5	276	2	F96583	hypothetical prote	507	35	36.5	696	2	E71510	probable oligopept
435	35	36.5	285	2	B83047	hypothetical prote	508	35	36.5	696	2	H81668	peptide ABC transp
436	35	36.5	290	2	A83816	hypothetical prote	509	35	36.5	698	2	C84836	hypothetical prote
437	35	36.5	292	2	D83580	probable transcrip	510	35	36.5	699	2	AG0371	probable acetyltra
438	35	36.5	292	2	T02320	hypothetical prote	511	35	36.5	704	2	T50303	hypothetical prote
439	35	36.5	296	2	AH2044	hypothetical prote	512	35	36.5	705	2	A71211	probable phosphori
440	35	36.5	309	2	H84657	hypothetical prote	513	35	36.5	724	2	T23322	hypothetical prote
441	35	36.5	311	2	T05561	probable uricase s	514	35	36.5	734	2	T24908	hypothetical prote
442	35	36.5	312	2	T34607	probable integral	515	35	36.5	744	2	S68115	catalase (EC 1.11.
443	35	36.5	314	2	F70505	probable trna delt	516	35	36.5	757	2	G71981	probable transcrip
444	35	36.5	316	2	AH1455	oligopeptide ABC t	517	35	36.5	759	1	C64345	replication licens
445	35	36.5	316	2	AH1091	oligopeptide ABC t	518	35	36.5	761	2	T41304	probable rna-bindi
446	35	36.5	318	2	S35145	carboxyphosphoen	519	35	36.5	762	1	A64420	replication licens
447	35	36.5	330	2	D95230	aspartate-ammonia	520	35	36.5	771	2	C75023	DNA polymerase I p
448	35	36.5	332	2	G98094	aspartate-ammonia	521	35	36.5	776	2	S45495	isp4 protein - fis
449	35	36.5	333	2	T26891	hypothetical prote	522	35	36.5	776	2	I48317	cyclin F - mouse
450	35	36.5	335	2	S21954	type II site-speci	523	35	36.5	779	2	C86335	hypothetical prote
451	35	36.5	344	1	GRYC57	sulfate transport	524	35	36.5	783	2	B84823	probable isomylas
452	35	36.5	349	2	E95858	conserved hypotet	525	35	36.5	821	2	T01181	hypothetical prote
453	35	36.5	358	1	WMBE38	infected cell prot	526	35	36.5	847	2	F83517	probable penicilli
454	35	36.5	359	2	S14028	type II site-speci	527	35	36.5	942	2	C81139	probable oxoglutar
455	35	36.5	372	2	B72472	probable electron	528	35	36.5	945	2	F83925	oxoglutarate dehyd
456	35	36.5	383	2	H82487	probable site-spec	529	35	36.5	956	2	H75536	2-oxoglutarate deh
457	35	36.5	389	2	T29488	hypothetical prote	530	35	36.5	956	2	S30834	hypothetical prote
458	35	36.5	397	2	B98293	mannonate dehydrat	531	35	36.5	962	2	T00262	hypothetical prote
459	35	36.5	397	2	AG2990	mannonate dehydrat	532	35	36.5	1014	1	DEBY	oxoglutarate dehyd
460	35	36.5	398	2	S40752	hypothetical prote	533	35	36.5	1022	2	T49683	probable oxoglutar
461	35	36.5	401	2	C83309	conserved hypotet	534	35	36.5	1023	2	T33116	probable phosphoe
462	35	36.5	424	2	T46197	hypothetical prote	535	35	36.5	1112	2	T30202	probable chitin sy
463	35	36.5	426	2	S53399	SFH1 protein - yea	536	35	36.5	1121	2	C87973	protein Y438C.12
464	35	36.5	430	2	S36233	leucoanthocyanidin	537	35	36.5	1127	2	T32404	hypothetical prote
465	35	36.5	455	2	F82345	conserved hypotet	538	35	36.5	1140	2	AE0180	probable membrane
466	35	36.5	458	2	S52594	chromosome segrega	539	35	36.5	1153	2	T26883	hypothetical prote
467	35	36.5	463	2	T16218	translation elonga	540	35	36.5	1162	2	C81297	proline dehydrogen

541	35	36.5	1214	2	G70953	probable suA prot	614	34	35.4	186	2	T42602	gene 59 protein -
542	35	36.5	1234	2	T30515	DNA-directed RNA p	615	34	35.4	191	2	E82447	pHnA protein VCA05
543	35	36.5	1235	2	C71120	probable DNA-direc	616	34	35.4	193	2	F84228	hypothetical prote
544	35	36.5	1260	2	A87046	hypothetical prote	617	34	35.4	195	2	JC2258	substrate protein
545	35	36.5	1290	2	A36466	1-phosphatidylinos	618	34	35.4	198	2	A05098	hypothetical prote
546	35	36.5	1290	2	A31117	1-phosphatidylinos	619	34	35.4	201	2	P70413	NADH2 dehydrogenas
547	35	36.5	1291	2	S00866	1-phosphatidylinos	620	34	35.4	203	2	A12445	hypothetical prote
548	35	36.5	1312	2	S68593	DNA-directed DNA p	621	34	35.4	206	2	S38626	glutathione transf
549	35	36.5	1317	2	B83346	probable non-ribos	622	34	35.4	220	2	JQ3075	chloramphenicol O-
550	35	36.5	1355	2	T22552	hypothetical prote	623	34	35.4	223	2	T37004	probable hydrolase
551	35	36.5	1373	1	A43291	collagen alpha 2(I	624	34	35.4	233	2	B97065	ribulose-5-phospha
552	35	36.5	1395	2	T00068	hypothetical prote	625	34	35.4	234	1	S76823	hypothetical prote
553	35	36.5	1404	2	T19277	hypothetical prote	626	34	35.4	245	2	C83377	conserved hypothet
554	35	36.5	1406	2	T17429	sag-pol polyprotei	627	34	35.4	246	2	C48492	3-deoxy-manno-octu
555	35	36.5	1449	2	A12017	hypothetical prote	628	34	35.4	252	2	T37633	succinate dehydrog
556	35	36.5	1503	2	T18266	cycloinulo-oligos	629	34	35.4	257	2	QJ0064	MERS protein - mou
557	35	36.5	1670	2	S71551	DNA-directed DNA p	630	34	35.4	261	2	J75556	hypothetical prote
558	35	36.5	1906	2	AD2443	hypothetical prote	631	34	35.4	272	2	AB1454	phage protein homo
559	35	36.5	2144	2	S71490	ashl protein - fru	632	34	35.4	272	2	AC1090	phage proteins hom
560	35	36.5	2241	2	T16064	hypothetical prote	633	34	35.4	273	1	QQRWTA	hypothetical 32K p
561	35	36.5	3147	2	T18674	hypothetical prote	634	34	35.4	273	2	S01245	hypothetical prote
562	35	36.5	3255	2	G81702	adherence factor T	635	34	35.4	273	2	S15609	hypothetical prote
563	35	36.5	4006	2	T09070	probable tenascin	636	34	35.4	273	2	T27068	hypothetical prote
564	35	36.5	5149	2	F83345	probable non-ribos	637	34	35.4	273	2	T20591	hypothetical prote
565	34.5	35.9	160	2	S72170	DNA-binding protei	638	34	35.4	273	2	B89599	protein F1085.9 [i
566	34.5	35.9	179	2	B84126	hypothetical prote	639	34	35.4	273	2	T16771	hypothetical prote
567	34.5	35.9	226	2	H82000	hypothetical prote	640	34	35.4	273	2	D88184	protein F18C5.7 [i
568	34.5	35.9	231	2	F95267	hypothetical prote	641	34	35.4	275	2	G71519	probable sam depen
569	34.5	35.9	318	2	G72310	hypothetical prote	642	34	35.4	278	2	T46458	hypothetical prote
570	34.5	35.9	339	2	T29405	conserved hypothet	643	34	35.4	280	2	F71829	siderophore-mediat
571	34.5	35.9	335	2	B70313	cytochrome-c perox	644	34	35.4	285	2	E64687	hypothetical prote
572	34.5	35.9	369	2	A69802	A/G-specific adeni	645	34	35.4	291	2	S74948	hypothetical prote
573	34.5	35.9	370	2	A95416	probable enoyl red	646	34	35.4	293	2	I40182	hydroxyquinol 1,2-
574	34.5	35.9	478	2	S16867	gene H5 protein -	647	34	35.4	297	2	A72130	hypothetical prote
575	34.5	35.9	489	2	S53005	histone-lysine N-m	648	34	35.4	298	1	A44878	protein kinase [EC
576	34.5	35.9	489	2	JC4324	lignostilbene alph	649	34	35.4	301	2	S38910	urate oxidase [EC
577	34.5	35.9	693	2	T00256	hypothetical prote	650	34	35.4	301	2	S73023	hypothetical prote
578	34.5	35.9	953	2	F96498	hypothetical prote	651	34	35.4	306	2	A75316	hypothetical prote
579	34.5	35.9	1102	2	T39943	hypothetical prote	652	34	35.4	311	2	E83804	protein-export mem
580	34.5	35.9	1106	2	T29496	hypothetical prote	653	34	35.4	314	2	T27721	hypothetical prote
581	34.5	35.9	1125	1	OYURCP	aperact receptor p	654	34	35.4	314	2	E87993	protein ZC34.4 [i
582	34.5	35.9	2183	2	T42764	coagulation factor	655	34	35.4	314	2	A87906	protein F59C6.1 [i
583	34.5	35.9	2211	1	KFB05	coagulation factor	656	34	35.4	314	2	G89343	protein F58B.3 [i
584	34.5	35.9	2907	2	A57278	fibrillin-2 precur	657	34	35.4	314	2	B88211	protein C28F5.5 [i
585	34	35.4	43	2	G69982	hypothetical prote	658	34	35.4	317	2	AF0973	lipopolysaccharide
586	34	35.4	44	2	I52861	vascular AIIa angi	659	34	35.4	317	2	A44156	lipopolysaccharide
587	34	35.4	48	2	PH0884	Ig kappa chain V r	660	34	35.4	322	2	B81001	heptosyltransferas
588	34	35.4	53	2	PH0881	Ig kappa chain V r	661	34	35.4	322	2	G82018	lipopolysaccharide
589	34	35.4	58	2	B82565	hypothetical prote	662	34	35.4	327	2	AC0156	probable cobalamin
590	34	35.4	106	2	T25617	hypothetical prote	663	34	35.4	329	2	AE2128	transcriptional regu
591	34	35.4	107	2	S36262	Ig lambda chain V	664	34	35.4	330	2	S70127	hypothetical prote
592	34	35.4	107	2	S36269	Ig lambda chain V	665	34	35.4	334	2	H95307	arginine deiminase
593	34	35.4	107	2	B98297	hypothetical 12.3K	666	34	35.4	334	2	E87448	arginine N-succiny
594	34	35.4	107	2	AB2986	conserved hypothet	667	34	35.4	340	2	T30121	hypothetical prote
595	34	35.4	108	2	A72704	hypothetical prote	668	34	35.4	343	2	T19597	hypothetical prote
596	34	35.4	110	2	B72661	hypothetical prote	669	34	35.4	343	2	T31588	hypothetical prote
597	34	35.4	111	2	PN0537	Ig kappa chain V r	670	34	35.4	343	2	F89402	protein Y94A7B.2 f
598	34	35.4	117	1	K4RBP2	Ig kappa chain pre	671	34	35.4	343	2	T18989	hypothetical prote
599	34	35.4	117	2	S42264	Ig kappa chain V r	672	34	35.4	343	2	E88061	protein ZK250.4 [i
600	34	35.4	117	2	S41809	Ig kappa chain V r	673	34	35.4	343	2	T33889	hypothetical prote
601	34	35.4	124	2	AC0803	probable bacteriop	674	34	35.4	346	2	AC0594	aldose 1-epimerase
602	34	35.4	129	2	B23986	Ig kappa chain pre	675	34	35.4	349	2	H95421	probable ABC trans
603	34	35.4	131	2	G84580	profilin 1 [import	676	34	35.4	350	2	F97254	probable translati
604	34	35.4	131	2	B85342	profilin 2 [import	677	34	35.4	352	2	I51282	TRP-1 - axolotl (f
605	34	35.4	135	2	H82531	preprotein translo	678	34	35.4	353	2	T35221	probable Atp/GTP b
606	34	35.4	144	2	A95285	hypothetical prote	679	34	35.4	356	2	T05119	leucoanthocyanidin
607	34	35.4	145	2	C96562	photosystem I subu	680	34	35.4	368	2	E82645	NAD (P)H-dependent
608	34	35.4	147	2	A36790	hypothetical prote	681	34	35.4	369	2	AG0625	conserved hypothet
609	34	35.4	148	2	A61237	E6 protein - human	682	34	35.4	373	2	T28978	hypothetical prote
610	34	35.4	148	2	S36573	E6 protein - human	683	34	35.4	383	2	D89975	conserved hypothet
611	34	35.4	161	2	T35260	hypothetical prote	684	34	35.4	384	2	T47213	probable trnaalose
612	34	35.4	163	2	G75424	bacterioferritin c	685	34	35.4	385	2	T11848	gibberellin 20-oxi
613	34	35.4	185	2	S48943	hypothetical prote	686	34	35.4	397	2	G84247	atrazine chlorohyd

687	34	35.4	400	2	150219	connexin 45.6 - ch	760	34	35.4	650	2	G84334	threonyl-tRNA synt
688	34	35.4	403	2	T47621	bZIP transcription	761	34	35.4	654	2	C87791	protein B0207.12 [
689	34	35.4	405	2	F89874	conserved hypotet	762	34	35.4	657	2	T42191	nikB protein - Esc
690	34	35.4	409	2	C82644	transcription regu	763	34	35.4	667	2	S60709	alpha 1,2 mannosid
691	34	35.4	412	2	F87460	hypothetical prote	764	34	35.4	673	2	T15551	hypothetical prote
692	34	35.4	415	2	AH2886	conserved hypotet	765	34	35.4	675	2	T23223	DNA topoisomerase
693	34	35.4	423	2	AG0451	malto porin [impor	766	34	35.4	685	2	A75055	probable DNA topoi
694	34	35.4	429	2	AH1009	sun protein [impor	767	34	35.4	686	1	E71106	translation IF2, G
695	34	35.4	430	2	S25977	ribosomal protein	768	34	35.4	693	2	D97122	helicase (EC 3.6.1
696	34	35.4	439	2	T32470	hypothetical prote	769	34	35.4	698	2	D64084	gelatinase B (EC 3
697	34	35.4	440	2	F83235	outer membrane por	770	34	35.4	708	2	JC4364	gelatinase B (EC 3
698	34	35.4	440	2	S11793	porin P precursor,	771	34	35.4	708	2	S62907	gelatinase B (EC 3
699	34	35.4	442	2	E97662	hypothetical prote	772	34	35.4	711	2	T48335	receptor like prot
700	34	35.4	443	2	T06809	protein kinase hom	773	34	35.4	711	2	A86424	unknown protein, 3
701	34	35.4	446	1	QRECL	lambda receptor pr	774	34	35.4	720	2	A36942	Fe(III)-pyochelin
702	34	35.4	446	2	C91256	maltose high-affin	775	34	35.4	727	2	E85896	probable peptidogl
703	34	35.4	446	2	C86096	maltose high-affin	776	34	35.4	728	2	T43632	serine/threonine p
704	34	35.4	449	2	C86496	hypothetical prote	777	34	35.4	729	2	A37191	glutamate-amonia
705	34	35.4	449	2	D72127	hypothetical prote	778	34	35.4	732	2	T19570	hypothetical prote
706	34	35.4	449	2	A81544	hypothetical prote	779	34	35.4	748	2	S59327	hypothetical prote
707	34	35.4	452	2	A60177	Lamb maltoporin pr	780	34	35.4	749	2	S64780	hypothetical prote
708	34	35.4	452	2	AF1014	malto porin precurs	781	34	35.4	750	1	A59145	methylmalonyl-CoA
709	34	35.4	455	2	I48724	zinc finger protei	782	34	35.4	756	2	T04187	subtilisin-like pr
710	34	35.4	462	2	E70955	hypothetical prote	783	34	35.4	759	2	D95129	pullulanase, proba
711	34	35.4	472	2	AC3534	glu/asp-tRNA amido	784	34	35.4	759	2	A98000	pullulanase (EC 3
712	34	35.4	477	2	S77373	hypothetical prote	785	34	35.4	762	2	A43355	cell surface prote
713	34	35.4	477	2	T32938	hypothetical prote	786	34	35.4	765	2	T04186	subtilisin-like pr
714	34	35.4	482	2	B31795	collagen alpha 1(X	787	34	35.4	769	2	D86335	T20H2, 6 protein -
715	34	35.4	482	2	T33651	hypothetical prote	788	34	35.4	769	2	H64525	transcription regu
716	34	35.4	483	2	G87493	conserved hypotet	789	34	35.4	770	2	A91052	penicillin binding
717	34	35.4	486	2	A13370	amino acid regulat	790	34	35.4	770	2	F65028	Penicillin-binding
718	34	35.4	486	2	D84912	hypothetical prote	791	34	35.4	774	2	B86492	hypothetical prote
719	34	35.4	496	2	A56040	protein-tyrosine k	792	34	35.4	774	2	H81540	hypothetical prote
720	34	35.4	508	1	NDBE83	exonuclease (EC 3	793	34	35.4	775	2	F98136	hypothetical prote
721	34	35.4	512	2	T13330	hypothetical prote	794	34	35.4	775	2	T04189	subtilisin-like pr
722	34	35.4	512	2	T04708	hypothetical prote	795	34	35.4	792	2	A83303	phenylalanyl-tRNA
723	34	35.4	513	2	S21535	nitrate reductase	796	34	35.4	792	2	T43630	serine/threonine p
724	34	35.4	513	2	T34689	probable iron-sulf	797	34	35.4	800	2	AE3151	glucoamylase [impo
725	34	35.4	517	2	AC2941	polysaccharuronase	798	34	35.4	807	2	E90523	leucyl-tRNA synthe
726	34	35.4	517	2	G98341	polysaccharuronase	799	34	35.4	811	2	E72003	hypothetical prote
727	34	35.4	518	2	JC7387	testis-abundant fi	800	34	35.4	823	2	H72282	hypothetical prote
728	34	35.4	520	2	D71368	cysteine-tRNA liga	801	34	35.4	823	2	S59141	translational initia
729	34	35.4	521	2	S69559	hypothetical prote	802	34	35.4	840	2	G83052	elongation factor
730	34	35.4	522	2	S33029	hypothetical prote	803	34	35.4	846	2	A96602	maltoose ABC transp
731	34	35.4	522	2	A81537	hypothetical prote	804	34	35.4	856	2	E75292	GDH family prote
732	34	35.4	523	2	T05934	protein kinase hom	805	34	35.4	867	1	GNLUSA	pol polyprotein -
733	34	35.4	526	1	GVNNSG	spike glycoprotein	806	34	35.4	869	2	A55384	transcription fact
734	34	35.4	527	2	D87318	conserved hypotet	807	34	35.4	878	2	C71305	leucine-tRNA ligas
735	34	35.4	536	2	T04234	calmodulin-binding	808	34	35.4	883	2	T23948	hypothetical prote
736	34	35.4	536	2	T52458	chlorophyll b synt	809	34	35.4	898	1	DJBPT4	DNA-directed DNA p
737	34	35.4	535	1	JX0225	cytochrome P450 Cy	810	34	35.4	913	2	AB2587	translation initia
738	34	35.4	547	2	D84546	probable importin,	811	34	35.4	913	2	E97369	translation initia
739	34	35.4	553	2	C75318	hypothetical prote	812	34	35.4	928	2	C97728	hypothetical prote
740	34	35.4	554	2	T51213	hypothetical prote	813	34	35.4	931	2	H86387	hypothetical prote
741	34	35.4	554	2	G75576	conserved hypotet	814	34	35.4	934	1	S42874	oxoglutarate dehyd
742	34	35.4	561	1	A41365	rhodopsin kinase (	815	34	35.4	935	2	A40694	cadherin-associate
743	34	35.4	567	2	T33650	hypothetical prote	816	34	35.4	936	2	H71728	2-oxoglutarate deh
744	34	35.4	571	1	XZADJ2	penton protein (II	817	34	35.4	937	2	A32879	oxoglutarate dehyd
745	34	35.4	571	1	XZADH5	penton protein (II	818	34	35.4	949	2	T16812	oxoglutarate dehyd
746	34	35.4	577	1	DNBYPA	polyadenylate-bind	819	34	35.4	960	2	T24221	hypothetical prote
747	34	35.4	585	2	F84346	archaeosine tRNA-r	820	34	35.4	967	2	T41672	probable b-zip tra
748	34	35.4	592	2	T50154	probable 6-phospho	821	34	35.4	980	2	H86197	hypothetical prote
749	34	35.4	598	2	T42592	protein-serine/thr	822	34	35.4	990	2	F88925	hypothetical prote
750	34	35.4	612	2	S53714	probable dinitrifi	823	34	35.4	996	2	G87687	hypothetical prote
751	34	35.4	616	1	VXZRNA	beta-1,2-glucan ex	824	34	35.4	1000	2	T21970	hypothetical prote
752	34	35.4	616	2	C75588	conserved hypotet	825	34	35.4	1004	2	AG3269	oxoglutarate dehyd
753	34	35.4	619	2	T00292	nikB protein - Esc	826	34	35.4	1009	2	T40412	2-oxoglutarate deh
754	34	35.4	619	2	T15193	hypothetical prote	827	34	35.4	1016	2	AC2900	oxoglutarate dehyd
755	34	35.4	626	2	C70801	hypothetical prote	828	34	35.4	1016	2	D97675	2-oxoglutarate deh
756	34	35.4	633	2	E86271	protein F16A14.4 [	829	34	35.4	1036	2	E84620	hypothetical prote
757	34	35.4	642	2	T40071	threonine-tRNA lig	830	34	35.4	1040	2	S50617	BSB1 protein - yea
758	34	35.4	643	2	S60710	alpha 1,2 mannosid	831	34	35.4	1049	1	A34286	NADPH-ferrihemopro
759	34	35.4	643	2	A41120	prostaglandin tran	832	34	35.4	1050	2	S54640	KCS1 protein - yea

833	34	35.4	1075	2	T27623	hypothetical prote	906	33	34.4	88	2	B83553	hypothetical prote
834	34	35.4	1080	2	T27622	hypothetical prote	907	33	34.4	89	2	A90339	hypothetical prote
835	34	35.4	1086	2	S74252	phosphorylase kina	908	33	34.4	90	2	G83316	cold-shock protein
836	34	35.4	1093	2	S74250	phosphorylase kina	909	33	34.4	102	2	F83231	hypothetical prote
837	34	35.4	1093	2	A31758	phosphorylase kina	910	33	34.4	104	2	AC0945	conserved hypotet
838	34	35.4	1093	2	B40793	phosphorylase kina	911	33	34.4	104	2	S40845	hypothetical 12.3K
839	34	35.4	1106	1	PFHUG3	platelet-derived g	912	33	34.4	104	2	D91232	hypothetical prote
840	34	35.4	1170	2	S30010	probable finger pr	913	33	34.4	104	2	C86079	hypothetical prote
841	34	35.4	1186	2	T19050	hypothetical prote	914	33	34.4	105	2	JE0384	NADH2 dehydrogenas
842	34	35.4	1206	2	S24407	formin isoform IV	915	33	34.4	107	2	S61363	prostaglandin-H E-
843	34	35.4	1299	2	AH2090	two-component hybr	916	33	34.4	107	2	D53285	Ig kappa chain V a
844	34	35.4	1307	2	G96711	unknown protein, 9	917	33	34.4	108	1	K1HULY	Ig kappa chain V-I
845	34	35.4	1350	2	T30341	zinc finger protei	918	33	34.4	108	2	S28241	NADH2 dehydrogenas
846	34	35.4	1367	2	T33819	hypothetical prote	919	33	34.4	111	2	H72501	hypothetical prote
847	34	35.4	1400	2	T52359	hypothetical prote	920	33	34.4	114	2	T50971	hypothetical prote
848	34	35.4	1702	2	S42459	DNA-directed DNA p	921	33	34.4	116	2	C75602	hypothetical prote
849	34	35.4	1747	2	T43162	vitellogenin - gvp	922	33	34.4	120	2	D95980	hypothetical prote
850	34	35.4	1806	1	CGHULE	collagen alpha 1(X	923	33	34.4	132	2	A55410	Ig light chain V r
851	34	35.4	1918	2	S43719	lactase (EC 3.2.1.	924	33	34.4	138	2	F91199	hypothetical prote
852	34	35.4	1920	2	S43721	lactase (EC 3.2.1.	925	33	34.4	149	2	AF2713	conserved hypotet
853	34	35.4	1926	2	S01169	beta-glycosidase c	926	33	34.4	156	2	A71189	hypothetical prote
854	34	35.4	1927	2	S01168	beta-glycosidase c	927	33	34.4	157	2	I53785	gene pp21 protein
855	34	35.4	1928	2	JS0610	beta-galactosidase	928	33	34.4	162	2	IS0017	MHC class I protei
856	34	35.4	2051	2	S34588	enoyl-lacyl-carrie	929	33	34.4	164	2	AG3189	hypothetical prote
857	34	35.4	2076	2	S15999	fatty-acyl-CoA syn	930	33	34.4	164	2	C97495	hypothetical prote
858	34	35.4	2121	2	T27406	hypothetical prote	931	33	34.4	169	2	F70799	hypothetical prote
859	34	35.4	2684	2	A96521	protein F21D18.22	932	33	34.4	171	2	S22194	hypothetical prote
860	34	35.4	3381	2	T42389	versican precursor	933	33	34.4	174	2	S53655	hydrogenase (EC 1.
861	34	35.4	3436	2	S55659	tegument protein 6	934	33	34.4	176	2	IS0016	MHC class I protei
862	34	35.4	3530	2	A59266	unconventional myo	935	33	34.4	177	2	AB2294	hypothetical prote
863	34	35.4	4485	2	T08044	dynein gamma heavy	936	33	34.4	178	2	F69804	hypothetical prote
864	34	35.4	4861	2	S71752	giant protein p619	937	33	34.4	179	2	T03215	translation elonga
865	34	35.4	13055	2	T16580	hypothetical prote	938	33	34.4	181	2	B82908	inorganic pyrophos
866	33.5	34.9	92	2	AC0993	probable exported	939	33	34.4	183	2	S31016	gene 71 protein -
867	33.5	34.9	188	2	S53077	hypothetical prote	940	33	34.4	187	2	C82613	conserved hypotet
868	33.5	34.9	197	2	S29574	Ig light chain - A	941	33	34.4	191	2	AG2710	cultivar specific
869	33.5	34.9	209	2	B87627	hypothetical prote	942	33	34.4	191	2	H97492	cultivar specific
870	33.5	34.9	281	2	C62431	hypothetical prote	943	33	34.4	194	2	A38203	proline-rich prote
871	33.5	34.9	306	1	W2WLEB	E2 protein - bovin	944	33	34.4	196	2	B48232	cysteine-rich exte
872	33.5	34.9	352	2	C83451	conserved hypotet	945	33	34.4	196	2	H72461	hypothetical prote
873	33.5	34.9	374	2	T33844	hypothetical prote	946	33	34.4	197	2	T36071	probable pantoate-
874	33.5	34.9	394	2	S20905	hypothetical prote	947	33	34.4	201	2	AC0627	conserved hypotet
875	33.5	34.9	403	2	A81464	translation initia	948	33	34.4	205	2	D82398	SM-20-related prot
876	33.5	34.9	441	2	F84560	purple acid phosph	949	33	34.4	206	2	B44941	glutathione transi
877	33.5	34.9	468	2	S47447	hypothetical prote	950	33	34.4	206	2	H47021	pectic enzyme secr
878	33.5	34.9	476	2	C72016	glycogen synthase	951	33	34.4	209	2	A48232	cysteine-rich exte
879	33.5	34.9	476	2	B86609	glycogen synthase	952	33	34.4	211	2	D95305	FixK-like regulato
880	33.5	34.9	476	2	H84291	hypothetical prote	953	33	34.4	214	2	AH2871	urease accessory p
881	33.5	34.9	492	2	AD0915	conserved hypotet	954	33	34.4	214	2	B97648	urease operon 23K
882	33.5	34.9	497	2	D65189	yigC protein - Esc	955	33	34.4	215	2	D91007	probable glutathio
883	33.5	34.9	497	2	C91225	probable oxidoredu	956	33	34.4	215	2	F85851	probable glutathio
884	33.5	34.9	497	2	A86072	probable oxidoredu	957	33	34.4	216	2	F84292	hypothetical prote
885	33.5	34.9	503	2	A40843	cytochrome P450 3	958	33	34.4	216	2	H87536	hypothetical prote
886	33.5	34.9	518	2	S60176	regulatory protein	959	33	34.4	216	2	AG1173	hypothetical prote
887	33.5	34.9	550	2	T40486	phosphoinositide-d	960	33	34.4	220	2	C96812	hypothetical prote
888	33.5	34.9	560	1	WFHUM	mullerian inhibiti	961	33	34.4	226	2	S43598	mers homolog R07B5
889	33.5	34.9	588	2	AB0591	succinate dehydrog	962	33	34.4	226	2	G95125	DNA repair protein
890	33.5	34.9	663	2	B65025	hypothetical prote	963	33	34.4	229	2	T05150	hypothetical prote
891	33.5	34.9	663	2	A91048	probable 2-compone	964	33	34.4	230	2	AF2237	hypothetical prote
892	33.5	34.9	679	2	E85892	hypothetical prote	965	33	34.4	231	2	A60468	venombin A (EC 3.4
893	33.5	34.9	750	2	S75632	sensory transducti	966	33	34.4	231	2	D97996	conserved hypotet
894	33.5	34.9	786	2	T49414	related to ahmpt p	967	33	34.4	231	2	B83031	chemotaxis protein
895	33.5	34.9	813	2	C97292	ATPases with chape	968	33	34.4	234	2	AB4253	nitrite hydratase
896	33.5	34.9	1163	2	JE0366	tight junction pro	969	33	34.4	235	1	B41326	hypothetical prote
897	33.5	34.9	1420	2	T37781	probable cytoskele	970	33	34.4	235	2	E72783	hypothetical prote
898	33.5	34.9	1473	2	T38791	probable ferredoxi	971	33	34.4	235	2	T37203	hypothetical prote
899	33.5	34.9	2104	2	D91286	hypothetical prote	972	33	34.4	236	2	I59323	apolipoprotein B m
900	33.5	34.9	2104	2	H86127	hypothetical prote	973	33	34.4	236	2	F84183	hypothetical prote
901	33.5	34.9	2111	2	T15390	hypothetical prote	974	33	34.4	241	2	S22457	Bronze-2 protein -
902	33.5	34.9	2224	1	KFHU5	coagulation factor	975	33	34.4	247	2	AD1381	carboxylesterase h
903	33	34.4	33	2	S53880	citQ protein - Lac	976	33	34.4	247	2	AE1750	carboxylesterase h
904	33	34.4	55	2	T16165	hypothetical prote	977	33	34.4	253	2	H70780	hypothetical prote
905	33	34.4	73	2	AE0258	hypothetical phage	978	33	34.4	255	2	T12074	urate oxidase (EC

979 33 34.4 255 2 C95903 probable transcript  
980 33 34.4 272 2 T09929 hypothetical prote  
981 33 34.4 273 2 E90124 putative myb like  
982 33 34.4 274 2 T31806 hypothetical prote  
983 33 34.4 275 2 S52278 hypothetical prote  
984 33 34.4 275 2 G83412 hypothetical prote  
985 33 34.4 279 2 I50014 MHC class I protei  
986 33 34.4 279 2 H83811 hypothetical prote  
987 33 34.4 282 2 H75537 phosphoprotein pho  
988 33 34.4 284 2 H75344 hypothetical prote  
989 33 34.4 286 2 C81681 dihydrodipicolinat  
990 33 34.4 286 2 F71524 dihydrodipicolinat  
991 33 34.4 286 2 A55349 regulatory protein  
992 33 34.4 288 2 G90368 hypothetical prote  
993 33 34.4 295 2 T35460 integral membrane  
994 33 34.4 296 2 A96278 peroxisomal acyl-c  
995 33 34.4 296 2 AB3006 conserved hypothet  
996 33 34.4 298 2 B87432 transcription regu  
997 33 34.4 298 2 T31518 hypothetical prote  
998 33 34.4 303 1 Z8BP22 gene 8 protein - p  
999 33 34.4 308 2 AC1645 N-acetylmuramoyl-L  
1000 33 34.4 309 1 A25776 urate oxidase (EC

## ALIGNMENTS

RESULT 1  
A39787  
teratocarcinoma-derived growth factor - human  
C;Species: Homo sapiens (man)  
C;Date: 20-Mar-1992 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
C;Accession: A39787  
R;Dono, R.; Montuori, N.; Rocchi, M.; De Ponti-Zilli, L.; Ciccodicola, A.; Persico, M.G.  
Am. J. Hum. Genet. 49, 555-565, 1991  
A;Title: Isolation and characterization of the CR1PTO autosomal gene and its X-linked re  
A;Reference number: A39787; UID:91353571; PMID:1882841  
A;Accession: A39787  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-188 <DON>  
A;Cross-references: UNIPROT:Q8TCC1; UNIPARC:UPI0000145185  
A;Note: the authors translated the codon GAC for residue 43 as Glu  
C;Superfamily: teratocarcinoma-derived growth factor 1; EGF homology

Query Match 100.0%; Score 96; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6.2e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWPOEPAIRPR 17  
Db 46 FRDSIWPOEPAIRPR 62

RESULT 2  
A30362  
teratocarcinoma-derived growth factor 1 - human  
N;Alternate names: CR1PTO protein  
C;Species: Homo sapiens (man)  
C;Date: 18-Oct-1989 #sequence\_revision 18-Oct-1989 #text\_change 09-Jul-2004  
C;Accession: B39787; A30362  
R;Dono, R.; Montuori, N.; Rocchi, M.; De Ponti-Zilli, L.; Ciccodicola, A.; Persico, M.G.  
Am. J. Hum. Genet. 49, 555-565, 1991  
A;Title: Isolation and characterization of the CR1PTO autosomal gene and its X-linked re  
A;Reference number: A39787; UID:91353571; PMID:1882841  
A;Accession: B39787  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-188 <DON>  
A;Cross-references: UNIPROT:P13385; UNIPARC:UPI000004966D; GB:M96955; GB:M37099; NID:g33  
R;Ciccodicola, A.; Dono, R.; Obici, S.; Simeone, A.; Zollo, M.; Persico, M.G.  
EMBO J. 8, 1987-1991, 1989  
A;Title: Molecular characterization of a gene of the 'EGF family' expressed in undiffer

A;Reference number: A30362; UID:90005403; PMID:2792079  
A;Accession: A30362  
A;Molecule type: mRNA  
A;Residues: 1-188 <CIC>  
A;Cross-references: UNIPARC:UPI000004966D; GB:X14253; NID:g30220; PIDN:CAA32467.1; PID:g  
C;Superfamily: teratocarcinoma-derived growth factor 1; EGF homology  
C;Keywords: growth factor  
F;78-106/Domain: EGF homology <EGF>

Query Match 100.0%; Score 96; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6.2e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWPOEPAIRPR 17  
Db 46 FRDSIWPOEPAIRPR 62

RESULT 3  
I49612  
teratocarcinoma-derived growth factor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I49612  
R;Dono, R.; Scalera, L.; Pacifico, F.; Simeone, A.; Persico, M.G.; Acampora, D.  
Development 118, 1157-1168, 1993  
A;Title: The murine cripto gene: expression during mesoderm induction and early heart m  
A;Reference number: I49612; UID:94094736; PMID:7916676  
A;Accession: I49612  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-171 <RES>  
A;Cross-references: UNIPROT:PS1865; UNIPARC:UPI0000029326; GB:M87321; NID:g402714; PIDN:  
C;Genetics:  
A;Gene: cripto  
C;Superfamily: teratocarcinoma-derived growth factor 1; EGF homology

Query Match 66.7%; Score 64; DB 2; Length 171;  
Best Local Similarity 68.8%; Pred. No. 0.0069;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RDDSINWQEPAIRPR 17  
Db 31 RDDSINWQEPAIRPR 46

RESULT 4  
B89027  
protein T08H10.1 [imported] - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 31-Dec-2004  
C;Accession: B89027  
R;anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A;Reference number: A75000; UID:99069613; PMID:9851916  
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C\_ele  
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A;Accession: B89027  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-496 <SPO>  
A;Cross-references: UNIPROT:Q22352; UNIPARC:UPI000017A66E; GB:chr\_V; PIDN:AAA97992.1; PT  
C;Genetics:  
A;Gene: T08H10.1  
A;Map position: 5  
C;Superfamily: Aldehyde reductase

Query Match 50.0%; Score 48; DB 2; Length 496;  
Best Local Similarity 40.0%; Pred. No. 7.9;  
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 RDDSINWQEPAIRPR 16









C;Accession: A25632  
R;Rafalski, J.A.  
Gene 43, 221-229, 1986  
A;Title: Structure of wheat gamma-gliadin genes.  
A;Reference number: A25632; MUID:86301876; PMID:3017812  
A;Accession: A25632  
A;Molecule type: DNA  
A;Residues: 1-291 <RAF>  
A;Cross-references: UNIPROT:P06659; UNIPARC:UPI000012B38C; GB:M13713; NID:g170707; PIDN:  
C;Comment: Gliadin, a glutamine and proline rich protein, is a major storage protein for  
C;Superfamily: gliadin  
C;Keywords: seed; storage protein  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-291/Product: gamma-gliadin B #status predicted <NAT>

Query Match 43.8%; Score 42; DB 1; Length 291;  
Best Local Similarity 50.0%; Pred. No. 39;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 7 WQOEPAIRP 16  
|||:|:|  
Db 31 WPOQPPFLQ 40

RESULT 14  
G87273  
oxidoreductase iron/ascorbate family CC0200 [imported] - Caulobacter crescentus  
C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C;Accession: G87273  
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
n, J.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: G87273  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-311 <STO>  
A;Cross-references: UNIPROT:Q9ABM7; UNIPARC:UPI00000C6F76; GB:AE005673; NID:g13421323; F  
C;Genetics:  
A;Gene: CC0200

Query Match 43.8%; Score 42; DB 2; Length 311;  
Best Local Similarity 50.0%; Pred. No. 42;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 DSIWQOEPAIR 15  
|||:|:|  
Db 116 DNVWPAEIPAFK 127

RESULT 15  
CSRZ  
catalase (EC 1.11.1.6) catA - rice  
C;Species: Oryza sativa (rice)  
C;Date: 31-Dec-1992 #sequence\_revision 02-Jun-2000 #text\_change 16-Jun-2000  
C;Accession: S70588; S20873; S19823  
R;Higo, K.; Higo, H.  
Plant Mol. Biol. 30, 505-521, 1996  
A;Title: Cloning and characterization of the rice CatA catalase gene, a homologue of the  
A;Reference number: S70588; MUID:96189265; PMID:8605302  
A;Accession: S70588  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-491 <HIG>  
A;Cross-references: UNIPARC:UPI000016DEAA; EMBL:D29966; NID:g1261857; PIDN:BAA06232.1; F  
R;Mori, H.; Higo, K.; Higo, H.; Minobe, Y.; Matsui, H.; Chiba, S.  
Plant Mol. Biol. 18, 973-976, 1992  
A;Title: Nucleotide and derived amino acid sequence of a catalase cDNA isolated from rice  
A;Reference number: S20873; MUID:92256818; PMID:1581574  
A;Accession: S20873

A;Molecule type: mRNA  
A;Residues: 1-112, 'Q', 114-447, 'AV', 450-491 <MOR>  
A;Cross-references: UNIPARC:UPI000016DEAC; EMBL:X61626; NID:g20191; PIDN:CAA43814.1; PID  
C;Genetics:  
A;Gene: cat A  
A;Introns: 5/3; 272/3; 472/3  
C;Superfamily: catalase  
C;Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
F;65,104,138/Active site: His, Ser, Asn #status predicted  
F;348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 43.8%; Score 42; DB 1; Length 491;  
Best Local Similarity 42.9%; Pred. No. 70;  
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 DSIWQOEPAIRP 16  
|||:|:|  
Db 289 DTKWPEDEVPLRP 302

RESULT 16  
F64889  
membrane protein maoC - Escherichia coli (strain K-12)  
C;Species: Escherichia coli  
C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
C;Accession: F64889  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
.A.; Rose, D.J.; Mau, B.; Shaoh, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: F64889  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-681 <BLAT>  
A;Cross-references: UNIPROT:P77455; UNIPARC:UPI000012EBAL; GB:AE000236; GB:U00096; NID:g  
A;Experimental source: strain K-12, substrain MGL55  
F;301-317/Domain: transmembrane #status predicted <TM01>  
F;588-604/Domain: transmembrane #status predicted <TM02>

Query Match 43.8%; Score 42; DB 2; Length 681;  
Best Local Similarity 55.6%; Pred. No. 1e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 DSIWQOE 11  
|||:|:|  
Db 120 DDTLWPEDE 128

RESULT 17  
S50163  
nitrate reductase (EC 1.7.99.4) large chain precursor, periplasmic - Thiophaera pantotro  
C;Species: Thiophaera pantotropha  
C;Date: 16-Feb-1995 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: S50163; S56135; S56128  
R;Berks, B.C.; Richardson, D.J.; Reilly, A.; Willis, A.C.; Ferguson, S.J.  
submitted to the EMBL Data Library, August 1994  
A;Description: The periplasmic nitrate reductase operon of Thiophaera pantotropha.  
A;Reference number: S50160  
A;Accession: S50163  
A;Molecule type: DNA  
A;Residues: 1-831 <BER>  
A;Cross-references: UNIPROT:Q56350; UNIPARC:UPI000012FD5F; EMBL:Z36773; NID:g600089; PIDN:  
R;Berks, B.C.; Richardson, D.J.; Reilly, A.; Willis, A.C.; Ferguson, S.J.  
Biochem. J. 309, 983-992, 1995  
A;Title: The napEDABC gene cluster encoding the periplasmic nitrate reductase system of  
A;Reference number: S56128; MUID:95366980; PMID:7639719  
A;Accession: S56135  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-31,42-89,154-203 <BEW>  
A;Cross-references: UNIPARC:UPI0000152014; UNIPARC:UPI0000174F6E; UNIPARC:UPI0000174F6F;  
A;Accession: S56128

A:Molecule type: protein  
A:Residues: 114-127, 'D', 129-130; 139-157; 317-339; 451-469; 591-600; 642-657; 675-694; 699-715  
A/Cross-references: UNIPARC:UPI0000174F70; UNIPARC:UPI0000174F71; UNIPARC:UPI0000174F72;  
F77  
C/Genetics:  
A/Gene: napA  
C/Superfamily: Alcaligenes eutrophus nitrate reductase A chain  
C/Keywords: blocked amino end; oxidoreductase  
F/1-31/Domain: signal sequence #status predicted <SIG>  
F/32-831/Product: nitrate reductase large chain #status predicted <MAT>

Query Match 43.8%; Score 42; DB 2; Length 831;  
Best Local Similarity 43.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 FRDSIWPOEPAIRP 16  
| | : || : |  
Db 572 FTDEWPFELIAMP 587

RESULT 18  
A47747  
tight junction protein ZO-1 - human  
C/Species: Homo sapiens (man)  
C/Date: 02-Jun-1994 #sequence\_revision 02-Jun-1994 #text\_change 09-Jul-2004  
R/Accession: A47747  
C/Willett, E.; Balda, M.S.; Fanning, A.S.; Jameson, B.; Van Itallie, C.; Anderson, J.M.  
Proc. Natl. Acad. Sci. U.S.A. 90, 7834-7838, 1993  
A/Title: The tight junction protein ZO-1 is homologous to the Drosophila discs-large tum  
A/Reference number: A47747; MUID:93361541; PMID:8395056  
A/Accession: A47747  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-1736 <WIL>  
A/Cross-references: UNIPROT:Q07157; UNIPARC:UPI000013C445; GB:L14837; NID:G292937; PIDN:  
C/Keywords: alternative splicing; membrane protein; phosphoprotein  
F/15-94/Domain: GLGF domain homology <GLG1>  
F/181-248/Domain: GLGF domain homology <GLG2>  
F/416-486/Domain: GLGF domain homology <GLG3>  
F/633-782/Domain: guanylate kinase homology <GKI>

Query Match 43.8%; Score 42; DB 2; Length 1736;  
Best Local Similarity 43.8%; Pred. No. 2.8e+02;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 RDSIWPOEPAIRP 17  
| : ||| : |  
Db 138 RSEKWPDRSASRR 153

RESULT 19  
S14275  
steroid 6beta-monooxygenase (EC 1.14.14.-) cytochrome P450 3A - dog  
N/Alternate names: steroid 6beta hydroxylase  
C/Species: Canis lupus familiaris (dog)  
C/Date: 23-Apr-1993 #sequence\_revision 23-Apr-1993 #text\_change 09-Jul-2004  
R/Accession: S14275  
C/Ciacio, P.J.; Graves, P.E.; Bourque, D.P.; Glinsmann-Gibson, B.; Halpert, J.R.  
Biochim. Biophys. Acta 1088, 319-322, 1991  
A/Title: cDNA and deduced amino acid sequences of a dog liver cytochrome P-450 of the III  
A/Reference number: S14275; MUID:91159488; PMID:2001406  
A/Accession: S14275  
A/Molecule type: mRNA  
A/Residues: 1-503 <CTA>  
A/Cross-references: UNIPROT:P24463; UNIPARC:UPI00001281CA; EMBL:X54915; NID:g909; PIDN:  
C/Genetics:  
A/Gene: CYP3A  
C/Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
C/Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo  
F/302-464/Domain: cytochrome P450 homology <P45>  
F/442/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 43.2%; Score 41.5; DB 2; Length 503;

A;Title: Genome structure of mycobacteriophage D29: Implications for phage evolution.  
A;Reference number: A72800; MUID:98300335; PMID:9636706  
A;Accession: F72803  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-223 <FOR>  
A;Cross-references: UNIPROT:O64226; UNIPARC:UPI0000138550; GB:AF022214; NID:g3172250; PID:  
C;Genetics:  
A;Gene: 32

Query Match 42.7%; Score 41; DB 2; Length 223;  
Best Local Similarity 35.7%; Pred. No. 42;  
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 DDSIWQEEPAIRP 16  
: ||| : : :  
Db 28 NQVWPAFDVLP 41

RESULT 23  
SNFF5K  
proteasome endopeptidase complex (EC 3.4.25.1) 35K chain - fruit fly (*Drosophila melanogaster*)  
N;Alternate names: 19S cylinder particle 35K chain; multicatalytic proteinase 35K chain;  
C;Species: *Drosophila melanogaster*  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C;Accession: S23450; S05507; A38761  
R;Frentzel, S.; Troxell, M.; Haas, C.; Pesold-Hurt, B.; Glaetzer, K.H.; Kloetzel, P.M.  
Eur. J. Biochem. 205, 1043-1051, 1992  
A;Title: Molecular characterization of the genomic regions of the *Drosophila* alpha-type  
A;Reference number: S23450; MUID:92249308; PMID:1374331  
A;Accession: S23450  
A;Molecule type: DNA  
A;Residues: 1-279 <FREN>  
A;Cross-references: UNIPROT:P12881; UNIPARC:UPI0000124852; EMBL:X62285; NID:g8387; PIDN:  
A;Experimental source: strain Canton S  
R;Haas, C.; Pesold-Hurt, B.; Mulchaup, G.; Beyreuther, K.; Kloetzel, P.M.  
EMBO J. 8, 2373-2379, 1989  
A;Title: The PROS-35 gene encodes the 35 kd protein subunit of *Drosophila melanogaster* P  
A;Reference number: S05507; MUID:90005444; PMID:2477245  
A;Accession: S05507  
A;Molecule type: mRNA  
A;Residues: 1-279 <HAA>  
A;Cross-references: UNIPARC:UPI0000124852; EMBL:X15497; NID:g8381; PIDN:CAA33520.1; PID:  
A;Accession: A38761  
A;Molecule type: protein  
A;Residues: 4-18,194-206 <HAA2>  
A;Cross-references: UNIPARC:UPI0000172D6B; UNIPARC:UPI0000172D6C  
C;Genetics:  
A;Gene: PROS-35  
A;Cross-references: FlyBase:FBgn0003151  
A;Map position: 89F-90A  
A;Introns: 1/3; 211/3  
C;Superfamily: multicatalytic endopeptidase complex chain C9  
C;Keywords: hydrolase; phosphoprotein; proteasome; protein degradation; proteinase

Query Match 42.7%; Score 41; DB 1; Length 279;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 DDSIWQEEPAIRP 17  
: ||| : : :  
Db 253 DRRSPPEEPAAGPR 267

RESULT 24  
C83883  
probable alpha-N-arabinofuranosidase (EC 3.2.1.55) BHI867 [similarity] - *Bacillus halodurans*  
C;Species: *Bacillus halodurans*  
C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C;Accession: C83883  
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai  
Nucleic Acids Res. 28, 4317-4331, 2000  
A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A;Reference number: A83650; MUID:20512582; PMID:11058132  
A;Accession: C83883  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-327 <STO>  
A;Cross-references: UNIPROT:Q9KB08; UNIPARC:UPI00000C3D11; GB:AP001513; GB:BA000004; NID:  
A;Experimental source: strain C-125  
C;Genetics:  
A;Gene: BHI867  
C;Superfamily: Streptomyces chartreusis alpha-L-arabinofuranosidase II  
C;Keywords: glycosidase; hydrolase

Query Match 42.7%; Score 41; DB 2; Length 327;  
Best Local Similarity 55.6%; Pred. No. 65;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 6 IWQPEPAI 14  
: ||| : : :  
Db 150 VWPQKDP 158

RESULT 25  
T50617  
hypothetical protein DKFZp762M15.1 - human (fragment)  
C;Species: *Homo sapiens* (man)  
C;Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 31-Dec-2004  
C;Accession: T50617  
R;Bloeker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Weil, B.; Wiemann, S.  
submitted to the Protein Sequence Database, June 2000  
A;Reference number: Z25143  
A;Accession: T50617  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-540 <AAA>  
A;Cross-references: UNIPROT:Q5NPX0; UNIPARC:UPI000006EF8A; EMBL:AL359587  
A;Experimental source: adult melanoma (Mewo cell line); clone DKFZp762M15  
C;Genetics:  
A;Note: DKFZp762M15.1  
C;Superfamily: 2-oxoglutarate dehydrogenase, E1 component; thiamin pyrophosphate-binding

Query Match 42.7%; Score 41; DB 2; Length 540;  
Best Local Similarity 77.8%; Pred. No. 11e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 DSIIWQEE 12  
: ||| : : :  
Db 479 DRIWSQEEP 487

RESULT 26  
C95854  
probable trehalose synthase protein [imported] - *Sinorhizobium meliloti* (strain 1021) mag  
C;Species: *Sinorhizobium meliloti*  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: C95854  
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A;Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo  
A;Reference number: A95842; MUID:21396508; PMID:11481431  
A;Accession: C95854  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-544 <KUR>  
A;Cross-references: UNIPROT:Q92X63; UNIPARC:UPI00000CB3F1; GB:AL591985; PIDN:CAC48499.1;  
A;Experimental source: strain 1021, megaplasmid pSymB  
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federici, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A;Reference number: A96039; MUID:21368234; PMID:11474104



A:Molecule type: DNA  
A:Residues: 1-721 <CON>  
A:Cross-references: UNIPROT:P40498; UNIPARC:UPI000013B420; GB:Z47047; EMBL:Z46728; NID:9  
C:Genetics:  
A:Gene: MIPS:YIL091C  
A:Cross-references: SGD:S0001353  
A:Map position: 9L  
C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPCC1827.01c

Query Match 42.7%; Score 41; DB 2; Length 721;  
Best Local Similarity 58.3%; Pred. No. 1.5e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FRDSTWQPEEP 12  
Db 358 FRDSTLPPKSKP 369  
|||||:|:|

RESULT 32  
T34797  
probable ATP binding protein - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004  
C:Accession: T34797  
R:Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, June 1998  
A:Reference number: Z21557  
A:Accession: T34797  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-756 <MUR>  
A:Cross-references: UNIPROT:O69900; UNIPARC:UPI00000DAC66; EMBL:AL023797; PIDN:CAA19404.  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC2E1.29

Query Match 42.7%; Score 41; DB 2; Length 756;  
Best Local Similarity 50.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 DSIWQPEEPAIRP 17  
Db 279 DSIWQPEEPAIRP 292  
|||||:|:|

RESULT 33  
G91015  
probable nitrate reductase 3 [imported] - Escherichia coli (strain O157:H7, substrain R1  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: G91015  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: G91015  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-828 <HAY>  
A:Cross-references: UNIPROT:Q8XE47; UNIPARC:UPI00000D02A5; GB:BA0000007; PIDN:BA036518.1  
A:Experimental source: strain O157:H7, substrain R1MD 0509952  
C:Genetics:  
A:Gene: EC83095  
C:Superfamily: Alcaligenes eutrophus nitrate reductase A chain

Query Match 42.7%; Score 41; DB 2; Length 828;  
Best Local Similarity 37.5%; Pred. No. 1.8e+02;  
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 FRDSTWQPEEPAIRP 16  
Db 569 FKTEVWPPELLAKP 584  
|:|:|:|:|

## RESULT 34

A85860  
probable nitrate reductase 3 napA [imported] - Escherichia coli (strain O157:H7, substra  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: A85860  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamamasis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: A85860  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-828 <STO>  
A:Cross-references: UNIPROT:Q8XE47; UNIPARC:UPI000015587B; GB:AE005174; NID:gl2516538; PJ  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: napA  
C:Superfamily: Alcaligenes eutrophus nitrate reductase A chain

Query Match 42.7%; Score 41; DB 2; Length 828;  
Best Local Similarity 37.5%; Pred. No. 1.8e+02;  
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 FRDSTWQPEEPAIRP 16  
Db 569 FKTEVWPPELLAKP 584  
|:|:|:|:|

## RESULT 35

T47438  
disease resistance protein homolog - Arabidopsis thaliana  
N:Alternate names: protein T18B22.30  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T47438  
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.  
submitted to the Protein Sequence Database, February 2000  
A:Reference number: Z24467  
A:Accession: T47438  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1214 <JOR>  
A:Cross-references: UNIPROT:Q9M1P1; UNIPARC:UPI000000C73F; EMBL:AL138652  
A:Experimental source: cultivar Columbia; BAC clone T18B22  
C:Genetics:  
A:Map position: 3  
A:Introns: 57/1; 238/2; 634/3; 736/3; 1206/1  
A:Note: T18B22.30

Query Match 42.7%; Score 41; DB 2; Length 1214;  
Best Local Similarity 57.1%; Pred. No. 2.8e+02;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 RDSIWQPEEPAIR 15  
Db 1115 RDSIWQPEEPAIR 1128  
|||||:|:|

## RESULT 36

A58612  
cytochrome P450 3A30 - mummichog (fragment)  
C:Species: Fundulus heteroclitus (mummichog)  
C:Date: 23-Sep-1997 #sequence\_revision 17-Oct-1997 #text\_change 09-Jul-2004  
C:Accession: A58612  
R:Celander, M.; Stegeman, J.J.  
Biochem. Biophys. Res. Commun. 236, 306-312, 1997  
A:Title: Isolation of a cytochrome P450 3A cDNA sequence (CYP3A30) from the marine teleost  
A:Reference number: A58612; MUID:97382427; PMID:9240431  
A:Accession: A58612

A;Molecule type: mRNA  
A;Residues: 1-124 <CEL>  
A;Cross-references: UNIPROT:Q9PVE8; UNIPARC:UPI0000174DAA  
A;Experimental source: liver  
C;Comment: This enzyme metabolizes steroids and a variety of non-steroids lipophilic org is involved in chemical carcinogenesis.  
C;Genetics:  
A;Gene: cyp3a  
C;Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
C;Keywords: heme

Query Match 42.2%; Score 40.5; DB 2; Length 124;  
Best Local Similarity 53.3%; Pred. No. 27;  
Matches 8; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 2 RDSIWPOEPAIRPR 16  
|||:|:|:|:  
Db 91 RDPEIWPEPE-AFPK 104

RESULT 37  
A82115  
sigma-54 dependent transcription activator VC2137 [imported] - Vibrio cholerae (strain N  
C;Species: Vibrio cholerae  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
A;Accession: A82115  
C;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A;Reference number: A82035; MUID:20406833; PMID:10952301  
A;Accession: A82115  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-488 <HBI>  
A;Cross-references: UNIPROT:Q9KQ66; UNIPARC:UPI00000C3201; GB:AE004286; GB:AE003852; NID  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:  
A;Gene: VC2137  
A;Map position: 1  
C;Superfamily: response regulator, NtrC type; response regulator homology; RNA polymerase

Query Match 42.2%; Score 40.5; DB 2; Length 488;  
Best Local Similarity 40.9%; Pred. No. 1.2e+02;  
Matches 9; Conservative 5; Mismatches 3; Indels 5; Gaps 1;

Qy 1 FRDD-----SIWPOEPAIRPR 17  
|||:|:|:|:|:|:  
Db 290 FREDLYVRLNVFPIEMPALRDR 311

RESULT 38  
AG0605  
oxygen-insensitive NADPH nitroreductase (EC 1.-.-.-) [imported] - Salmonella enterica su  
C;Species: Salmonella enterica subsp. enterica serovar Typhi  
A;Note: this species has also been called Salmonella typhi  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
A;Accession: AG0605  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, P.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A;Reference number: AB0502; MUID:21534947; PMID:11677608  
A;Accession: AG0605  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-240 <PAR>  
A;Cross-references: UNIPARC:UPI000005A0B4; GB:AL513382; PIDN:CAD05313.1; PID:G16502077;  
C;Genetics:  
A;Gene: nfsA

C;Superfamily: NADPH-flavin oxidoreductase homolog  
C;Keywords: oxidoreductase

Query Match 41.7%; Score 40; DB 2; Length 240;  
Best Local Similarity 45.5%; Pred. No. 66;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 7 WPOEPAIRPR 17  
|||:|:|:|:  
Db 159 WPADNPDLKPR 169

RESULT 39  
I80318  
drug activity modulator A - Escherichia coli (strain K-12)  
N;Contains: aromatic nitrate reductase (NADPH) (EC 1.6.6.-) oxygen-insensitive  
C;Species: Escherichia coli  
C;Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 09-Jul-2004  
C;Accession: I80318; C64823; S04774  
R;Chatterjee, P.K.; Sternberg, N.L.  
Proc. Natl. Acad. Sci. U.S.A. 92, 8950-8954, 1995  
A;Title: A general genetic approach in Escherichia coli for determining the mechanism(s)  
A;Reference number: I59418; MUID:96004656; PMID:7568050  
A;Accession: I80318  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-240 <RES>  
A;Cross-references: UNIPROT:P17117; UNIPARC:UPI00000480E3; EMBL:U18655; NID:G609323; PID  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
.A.; Rose, D.J.; Mau, B.; Shaio, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: C64823  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Residues: 1-240 <BLAT>  
A;Cross-references: UNIPARC:UPI00000480E3; GB:AE000187; GB:U000096; NID:g1787070; PIDN:AA  
A;Experimental source: strain K-12, substrain MG1655  
R;Kang, W.K.; Icho, T.; Isono, S.; Kitakawa, M.; Isono, K.  
Mol. Gen. Genet. 217, 281-288, 1989  
A;Title: Characterization of the gene rimK responsible for the addition of glutamic acid  
A;Reference number: S04774; MUID:89364710; PMID:2570347  
A;Accession: S04774  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 134-154, 'CA', 157, 'AGH', 161-186, 'RHRSWTSWNRNITSPVAIIAGIIPGAILIAEQSLKKAHLFWI  
C;Cross-references: UNIPARC:UPI000016F48A; GB:X15859; NID:g42745; PIDN:CRA33867.1; PID:g  
C;Genetics:  
A;Gene: mdaA; mda18  
A;Map position: 18.7-19.0 min  
C;Superfamily: NADPH-flavin oxidoreductase homolog  
C;Keywords: flavoprotein; FMN; oxidoreductase

Query Match 41.7%; Score 40; DB 2; Length 240;  
Best Local Similarity 45.5%; Pred. No. 66;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 7 WPOEPAIRPR 17  
|||:|:|:|:  
Db 159 WPADNPDLKPR 169

RESULT 40  
C90745  
modulator of drug activity A [imported] - Escherichia coli (strain O157:H7, substrain RI  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
A;Accession: C90745  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: C90745

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-240 <HAY>

A:Cross-references: UNIPROT:Q8X6S1; UNIPARC:UPI00000D0B32; GB:BA0000007; PIDN:BAB34354.1;

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: EC60931

C:Superfamily: NADPH-flavin oxidoreductase homolog

Query Match 41.7%; Score 40; DB 2; Length 240;

Best Local Similarity 45.5%; Pred. No. 66;

Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 WPOEPAIRPR 17

|||:|:|

Db 159 WPADNFDLKP 169

RESULT 41

G85595

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-240 <STO>

A:Cross-references: UNIPROT:Q8X6S1; UNIPARC:UPI00000D0B32; GB:AE005174; NID:g12513866; E

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: mdaA

C:Superfamily: NADPH-flavin oxidoreductase homolog

Query Match 41.7%; Score 40; DB 2; Length 240;

Best Local Similarity 45.5%; Pred. No. 66;

Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 WPOEPAIRPR 17

|||:|:|

Db 159 WPADNFDLKP 169

RESULT 42

E72703

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-249 <KAW>

A:Cross-references: UNIPROT:Q9YD68; UNIPARC:UPI000005DDBA; DDBJ:AP000060; NID:g5104188;

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1044

C:Superfamily: NADPH-flavin oxidoreductase homolog

Query Match 41.7%; Score 40; DB 2; Length 249;

Best Local Similarity 54.5%; Pred. No. 69;

Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 WPOEPAIRPR 17

|||:|:|

Db 159 WPADNFDLKP 169

RESULT 43

S74432

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-312 <KAN>

A:Cross-references: UNIPROT:P72585; UNIPARC:UPI00000C0998; EMBL:D90899; GB:AB001339; NID

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

A:Superfamily: Escherichia coli probable UDPglucose 4-epimerase

Query Match 41.7%; Score 40; DB 1; Length 312;

Best Local Similarity 41.2%; Pred. No. 89;

Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 FRDSDIWPOEPAIRPR 17

|||:|:|

Db 260 FECDIITWTDQNGQPR 276

RESULT 44

AG2956

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-324 <KUR>

A:Cross-references: UNIPROT:Q8UAW4; UNIPARC:UPI00000D212F; GB:AE008689; PIDN:AAL44069.1;

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu3253

A:Map position: linear chromosome

C:Superfamily: conserved hypothetical protein HI1028

Query Match 41.7%; Score 40; DB 2; Length 324;

Best Local Similarity 53.3%; Pred. No. 92;

Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRDSDIWPOEPAIR 15

|||:|:|

Db 40 FHTDSVWAAEIAKR 54

RESULT 45

G98326



c4-dicarboxylate-binding periplasmic protein VC1779 [imported] - Agrobacterium tumefaciens  
C;Species: Agrobacterium tumefaciens  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C;Accession: G98326  
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurolo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A;Reference number: A97359; MUID:21608551; PMID:11743194  
A;Accession: G98326  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-324 <KUR>  
A;Cross-references: UNIPROT:Q8UAW4; UNIPARC:UPI00000D212F; GB:AE007870; PIDN:AAK90137.1;  
C;Genetics:  
A;Gene: AGR\_L\_3125  
A;Map position: linear chromosome  
C;Superfamily: conserved hypothetical protein HI1028

Query Match 41.7%; Score 40; DB 2; Length 324;  
Best Local Similarity 53.3%; Pred. No. 92;  
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 FRDSIWPOEPAIR 15  
| | | | |  
Db 40 FHTDSVWAEIACK 54

RESULT 46  
C97268  
uncharacterized protein, YceG B. subtilis homolog [imported] - Clostridium acetobutylicum  
C;Species: Clostridium acetobutylicum  
C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: C97268  
R;Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R. J. Bacteriol. 183, 4823-4836, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: C97268  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-328 <KUR>  
A;Cross-references: UNIPROT:Q9TEW3; UNIPARC:UPI00000C36C7; GB:AE001437; PIDN:AAK80934.1;  
A;Experimental source: Clostridium acetobutylicum ATCC824  
C;Genetics:  
A;Gene: CAC2993

Query Match 41.7%; Score 40; DB 2; Length 328;  
Best Local Similarity 58.3%; Pred. No. 94;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 SIWQEEPAIRP 16  
| | | | |  
Db 58 NIWSKEAIRP 69

RESULT 47  
T01256  
probable anthocyanidin synthase [imported] - Arabidopsis thaliana  
N;Alternate names: SRG1 protein homolog F16M14.17  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C;Accession: T01256; F84802  
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, July 1998  
A;Description: Arabidopsis thaliana chromosome II BAC F16M14 genomic sequence.  
A;Reference number: Z14213  
A;Accession: T01256  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-353 <ROU>  
A;Cross-references: UNIPROT:O80449; UNIPARC:UPI0000048612; EMBL:AC003028; NID:g3335356;

A;Experimental source: cultivar Columbia  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.W.; Venter, L. Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: F84802  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-353 <STO>  
A;Cross-references: UNIPARC:UPI0000048612; GB:AE002093; NID:g3335372; PIDN:AACT27173.1; G  
C;Genetics:  
A;Gene: F16M14.17; At2g38240  
A;Map position: 2; 271/3  
A;Introns: 160/2; 271/3  
C;Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 41.7%; Score 40; DB 2; Length 353;  
Best Local Similarity 50.0%; Pred. No. 1e+02;  
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 RDSIWPOEPAIR 15  
| | | | |  
Db 147 RNPSKWPSQPPKIR 160

RESULT 48  
A55589  
lysine-tRNA ligase (EC 6.1.1.6) - Thermus aquaticus  
N;Alternate names: lysyl-tRNA synthetase  
C;Species: Thermus aquaticus  
C;Date: 03-Mar-1995 #sequence\_revision 03-Mar-1995 #text\_change 03-Jun-2002  
C;Accession: A55589  
R;Chen, J.; Brevet, A.; Lapadat-Tapolsky, M.; Blanquet, S.; Plateau, P. J. Bacteriol. 176, 2699-2705, 1994  
A;Title: Properties of the lysyl-tRNA synthetase gene and product from the extreme thermophile Thermus aquaticus  
A;Reference number: A55589; MUID:94222851; PMID:8169220  
A;Accession: A55589  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-492 <CHE>  
A;Cross-references: UNIPARC:UPI0000136547; GB:X70708; NID:g486664; PIDN:CAA50039.1; PID:  
C;Genetics:  
A;Gene: lysS  
C;Superfamily: lysine-tRNA ligase  
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 41.7%; Score 40; DB 2; Length 492;  
Best Local Similarity 50.0%; Pred. No. 1e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 4 DSIWQEEPAIRPR 17  
| | | | |  
Db 46 ESEWPEEEVAVAGR 59

RESULT 49  
T36294  
dihydroxy-acid dehydratase - Streptomyces coelicolor  
C;Species: Streptomyces coelicolor  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 05-Oct-2004  
C;Accession: T36294  
R;Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, May 1998  
A;Reference number: Z21603  
A;Accession: T36294  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-617 <SEE>  
A;Cross-references: UNIPROT:O69198; UNIPARC:UPI000012D58E; EMBL:AL049819; PIDN:CAB42671.  
A;Experimental source: strain A3(2)  
C;Genetics:



A:Gene: SCOEDB:SCE7.12c  
C:Superfamily: Dihydroxy-acid dehydratase/phosphogluconate dehydratase

Query Match 41.7%; Score 40; DB 2; Length 617;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 DDSIWPOEPAI 14  
|:|:|:|:|:|:  
Db 448 DESIWTPEGPAV 459

RESULT 50  
T16607  
hypothetical protein K10B2.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T16607  
R:Miller, N.  
submitted to the EMBL Data Library, June 1995  
A:Description: The sequence of C. elegans cosmid K10B2.  
A:Reference number: Z18545  
A:Accession: T16607  
A>Status: preliminary; translated from GE/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-701 <MIL>  
A:Cross-references: UNIPARC:UPI000017BABA; EMBL:U28730; NID:g860694; PID:g860695; PIDN:A  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:K10B2.1  
A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3

Query Match 41.7%; Score 40; DB 2; Length 701;  
Best Local Similarity 87.5%; Pred. No. 2.2e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 EEPALRPR 17  
|:|:|:|:|:|:  
Db 587 EEPALRPR 594

Search completed: February 3, 2006, 18:56:48  
Job time : 44 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM protein - protein search, using sw model  
Run on: February 3, 2006, 18:39:32 ; Search time 316.5 Seconds  
(without alignments)  
23.600 Million cell updates/sec

Title: US-10-693-538-1-copy\_46\_62  
Perfect score: 96  
Sequence: 1 FRDSIWQEPQAIRPR 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : A Geneseq\_21.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	96	100.0	17 7 ADD26060
2	96	100.0	129 5 ABB77111
3	96	100.0	129 5 ABB77124
4	96	100.0	129 5 ABB77104
5	96	100.0	129 5 ABB77117
6	96	100.0	139 5 ABB77103
7	96	100.0	139 5 ABB77123
8	96	100.0	139 5 ABB77110
9	96	100.0	139 5 ABB77116
10	96	100.0	173 8 ADS88698
11	96	100.0	174 2 AAR13326
12	96	100.0	174 2 AAW32107
13	96	100.0	174 8 ADO05065
14	96	100.0	183 8 ADS88699
15	96	100.0	184 8 ADO28629
16	96	100.0	187 6 AAE36467
17	96	100.0	188 2 AAR22548
18	96	100.0	188 2 AAR22547
19	96	100.0	188 2 AAW25667
20	96	100.0	188 2 AAW32108
21	96	100.0	188 2 AAW19980
22	96	100.0	188 2 AAW29735
23	96	100.0	188 2 AAW87631
24	96	100.0	188 2 AAW87630

25	96	100.0	188	5	ABB77102	Human Cri
26	96	100.0	188	5	ABB77101	Human Cri
27	96	100.0	188	5	ABB77122	Human Cri
28	96	100.0	188	5	ABB77109	Human Cri
29	96	100.0	188	5	AAO14637	Human Cri
30	96	100.0	188	5	AAO14638	Human Cri
31	96	100.0	188	5	AAO14636	Human Cri
32	96	100.0	188	5	AAO14727	Human var
33	96	100.0	188	6	ABP97176	Tumour-as
34	96	100.0	188	6	ABP58132	Human Cri
35	96	100.0	188	6	ABP58131	Human Cri
36	96	100.0	188	7	ADC78799	Human PRO
37	96	100.0	188	7	ADC78843	Human PRO
38	96	100.0	188	7	ADD26059	CR-2 crip
39	96	100.0	188	7	ADD26058	CR-1 crip
40	96	100.0	188	8	ADI82176	Human ter
41	96	100.0	188	8	ADR70486	Human ter
42	96	100.0	188	8	ADS88697	Amino aci
43	96	100.0	188	8	ADT79198	Human Cri
44	96	100.0	188	8	ADT79200	Human Cri
45	96	100.0	188	8	ADT79197	Human Cri
46	96	100.0	188	8	ADT79290	Human Cri
47	96	100.0	188	9	ADY85961	Human Cri
48	96	100.0	188	9	ADZ42238	Human Cri
49	96	100.0	360	5	ABB77114	Human Cri
50	96	100.0	360	5	ABB77107	Human Cri
51	96	100.0	360	5	ABB77120	Human Cri
52	96	100.0	360	5	ABB77127	Human Cri
53	96	100.0	367	5	ABB77106	Human Cri
54	96	100.0	367	5	ABB77119	Human Cri
55	96	100.0	367	5	ABB77113	Human Cri
56	96	100.0	367	5	ABB77126	Human Cri
57	64	66.7	96	8	ADS88696	Amino aci
58	64	66.7	156	8	ADS88691	Amino aci
59	64	66.7	156	8	ADS88692	Amino aci
60	64	66.7	171	8	ADS88690	Amino aci
61	64	66.7	171	9	ADY85960	Murine Cr
62	64	66.7	171	9	ADZ42246	Mouse Cri
63	54	56.2	9	5	AAO14712	Human Cri
64	54	56.2	9	8	ADT79274	Human Cri
65	48	50.0	496	8	ADN23652	Bacterial
66	48	50.0	606	4	ABBS7811	Drosophil
67	48	50.0	606	4	ABBS7812	Drosophil
68	48	50.0	1081	3	AAV93499	Amino aci
69	48	50.0	1081	5	ABG93377	S. cerevi
70	47	49.0	18	5	AAO14644	Human Cri
71	47	49.0	18	8	ADT79206	Human Cri
72	47	49.0	137	8	ABO55359	Human gen
73	47	49.0	216	8	ADX92990	Plant ful
74	47	49.0	348	8	ADT57265	Plant pol
75	46	47.9	252	7	ABO69627	Pseudomon
76	46	47.9	988	8	ADK98519	S Bulboca
77	46	47.9	988	8	ADK98517	S Bulboca
78	46	47.9	988	8	ADK98525	S Bulboca
79	46	47.9	1003	8	ADK98521	S Bulboca
80	45.5	47.4	74	6	ABM44956	Propionib
81	45.5	47.4	74	6	ABM44956	Propionib
82	45	46.9	104	6	ABM44956	Propionib
83	45	46.9	104	6	ABM44956	Propionib
84	45	46.9	184	7	ADP13889	Human pol
85	45	46.9	212	4	AAM42120	Human end
86	45	46.9	415	9	AEB36512	L. pneumo
87	45	46.9	415	9	AEB42298	L. pneumo
88	45	46.9	448	7	ABO81484	Pseudomon
89	45	46.9	490	4	AAM40334	Human pol
90	45	46.9	490	5	ABBS84260	Human kat
91	45	46.9	490	5	AAU74352	Human cyt
92	45	46.9	490	7	ADM05650	Human pro
93	45	46.9	562	4	ABG15573	Novel hum
94	45	46.9	766	3	AAV93498	Amino aci
95	45	46.9	945	8	ADH51542	Plant inf
96	45	46.9	1572	4	ABBS2868	Drosophil
97	44.5	46.4	906	8	ADN23873	Bacterial

98	44	45.8	55	4	ABO3527	Human mus	171	42	43.8	945	8	ADH51545	Plant inf
99	44	45.8	55	6	ABU12821	Novel hum	172	42	43.8	945	8	ADH51546	Plant inf
100	44	45.8	55	8	ADJ28847	Human mus	173	42	43.8	970	7	ADF17765	Solanum b
101	44	45.8	115	4	AAU14145	Human nov	174	42	43.8	970	8	ADH51537	S bulboca
102	44	45.8	349	9	ADM18198	E grandis	175	42	43.8	970	9	ADX97881	Solanum b
103	44	45.8	349	9	ADM18624	Eucalyptu	176	42	43.8	979	7	ADF17766	Solanum b
104	44	45.8	608	6	ABU20989	Protein e	177	42	43.8	979	8	ADH51538	Plant inf
105	44	45.8	1040	7	ADF17768	Solanum b	178	42	43.8	979	8	ADH51541	Plant inf
106	44	45.8	1040	8	ADH51540	Plant inf	179	42	43.8	1096	4	ABG08906	Novel hum
107	44	45.8	1312	5	ABB77984	Amino aci	180	42	43.8	1427	6	ABU21601	Protein e
108	44	45.8	1312	6	ABR53887	Protein s	181	42	43.8	1736	5	AAU84308	Human end
109	44	45.8	1312	7	ADK63602	Disease t	182	42	43.8	1736	8	ADN04661	Antipsori
110	44	45.8	1312	8	ADN19297	Bacterial	183	42	43.8	1771	8	ADR66853	Human pro
111	43	44.8	99	5	ABP34477	Human ORF	184	42	43.8	1771	8	ADR66292	Human pro
112	43	44.8	167	4	AG733980	Human col	185	42	43.8	1906	4	ABB62803	Drosophil
113	43	44.8	195	8	ADY09143	Plant ful	186	41.5	43.2	503	8	ADO47244	Dog cytoc
114	43	44.8	335	6	ABU23173	Protein e	187	41.5	43.2	944	5	ABG93384	C. albica
115	43	44.8	416	8	ABO59700	Human gen	188	41	42.7	59	4	AAU60218	Propionib
116	43	44.8	620	7	ADM04773	Human pro	189	41	42.7	59	6	ABM56737	Propionib
117	43	44.8	663	7	ADM04807	Human pro	190	41	42.7	104	7	ADM05723	Human pro
118	43	44.8	707	4	ABB68532	Drosophil	191	41	42.7	152	4	AAU43463	Propionib
119	43	44.8	776	8	ADO58207	Mutant JD	192	41	42.7	152	4	ABM33982	Propionib
120	43	44.8	776	8	ADN75157	JDF-3 pol	193	41	42.7	169	4	ABM11443	Human LMW
121	43	44.8	776	8	ADQ26688	Mutant JD	194	41	42.7	197	6	ABU25563	Protein e
122	43	44.8	776	8	ADT07163	Pfu mutan	195	41	42.7	223	5	ADK35509	Novel hum
123	43	44.8	841	9	ADY14116	Human src	196	41	42.7	227	4	ABG03220	Novel hum
124	43	44.8	841	9	ADY70338	Human bet	197	41	42.7	227	4	ABG22343	Novel hum
125	43	44.8	1084	5	AAE24146	Human kin	198	41	42.7	235	7	ABM86320	Rice abio
126	43	44.8	1113	5	AAE17836	Herpes si	199	41	42.7	256	5	ABP69759	Human pol
127	43	44.8	1113	7	ADG74979	Human her	200	41	42.7	279	4	ABB61086	Drosophil
128	43	44.8	1114	2	AAW72205	HSV-2 str	201	41	42.7	301	4	ABB70654	Drosophil
129	43	44.8	1114	2	AAW72096	HSV-2 str	202	41	42.7	448	4	ABG26363	Novel hum
130	43	44.8	1114	7	ADG75130	Human her	203	41	42.7	474	6	ABU33592	Protein e
131	43	44.8	1231	9	ADX98205	Lysine de	204	41	42.7	474	9	ABE41070	L. pneumo
132	43	44.8	1243	6	AAE34865	Human kin	205	41	42.7	476	9	ABE37760	L. pneumo
133	43	44.8	1243	6	ABP71620	Human WNK	206	41	42.7	522	8	ADP69071	Human NOV
134	43	44.8	1513	4	AAU03531	Human pro	207	41	42.7	522	8	ADP69067	Human NOV
135	42.5	44.3	921	4	ABB63852	Drosophil	208	41	42.7	522	8	ADP69063	Human NOV
136	42	43.8	20	8	ADH15855	Gliadin r	209	41	42.7	522	8	ADP69069	Human NOV
137	42	43.8	20	8	ADH15855	Gliadin r	210	41	42.7	523	8	ADP69065	Human NOV
138	42	43.8	103	4	ABB16235	Human ner	211	41	42.7	541	5	ABG61769	Novel sec
139	42	43.8	107	9	AEA40477	Anti-VEGF	212	41	42.7	541	6	ADA55210	Human pro
140	42	43.8	107	9	AEA40480	Anti-VEGF	213	41	42.7	541	8	ADK52111	Human ato
141	42	43.8	107	5	ABB89546	Human pol	214	41	42.7	589	6	ABM69298	Photorhab
142	42	43.8	192	8	ADJ57873	Human imm	215	41	42.7	590	7	ABG66465	Klebsiell
143	42	43.8	219	7	ABM85965	Rice abio	216	41	42.7	639	6	ABU34919	Protein e
144	42	43.8	236	8	ADQ14475	Mouse ant	217	41	42.7	639	6	ABU36770	Protein e
145	42	43.8	255	7	ABO63928	Klebsiell	218	41	42.7	642	6	ABU35836	Protein e
146	42	43.8	308	7	ABM74152	DNA clone	219	41	42.7	717	5	ABG79520	Human tum
147	42	43.8	311	8	ADS27739	Bacterial	220	41	42.7	774	8	ADO58205	Mutant Ve
148	42	43.8	349	7	ABO80709	Pseudomon	221	41	42.7	774	8	ADQ26686	Mutant Ve
149	42	43.8	420	8	ABO58470	Human gen	222	41	42.7	774	8	ADT07161	Pfu mutan
150	42	43.8	491	8	ADQ37161	Cell prol	223	41	42.7	889	8	ADQ15122	Human can
151	42	43.8	491	8	ADQ15745	Rice stre	224	41	42.7	971	4	ABG14701	Novel hum
152	42	43.8	554	6	ABR40706	Triticum	225	41	42.7	971	7	ADF60445	Human con
153	42	43.8	615	4	AAE93480	Human pro	226	41	42.7	971	7	ADF60349	Human con
154	42	43.8	615	4	AAU28187	Novel hum	227	41	42.7	972	8	ADH51544	Plant inf
155	42	43.8	673	4	ABG26099	Novel hum	228	41	42.7	992	7	ADF17767	Plant inf
156	42	43.8	681	8	ADU03333	E coli Ma	229	41	42.7	992	8	ADH51539	Plant inf
157	42	43.8	773	8	ADO58201	Mutant Tg	230	41	42.7	992	8	ADH51543	Plant inf
158	42	43.8	773	8	ADN75163	Tgo polym	231	41	42.7	1185	7	ABM84878	Rice abio
159	42	43.8	773	8	ADQ26682	Mutant TG	232	41	42.7	1233	8	ADN20989	Bacterial
160	42	43.8	773	8	ADT07157	Pfu mutan	233	40.5	42.2	367	8	ADX95734	Plant ful
161	42	43.8	774	8	ADO58203	Mutant KO	234	40.5	42.2	428	8	ADX97312	Plant ful
162	42	43.8	774	8	ADN75151	KOD polym	235	40.5	42.2	428	8	ADX90673	Plant ful
163	42	43.8	774	8	ADQ26684	Mutant KO	236	40	41.7	65	8	ADP29563	Human sec
164	42	43.8	774	8	ADT07159	Pfu mutan	237	40	41.7	75	4	AAU48136	Propionib
165	42	43.8	775	8	ADO58199	Mutant De	238	40	41.7	75	6	ABM44655	Propionib
166	42	43.8	775	8	ADQ26680	Mutant De	239	40	41.7	78	4	AAU41559	Propionib
167	42	43.8	775	8	ADT07155	Pfu mutan	240	40	41.7	78	6	ABM38078	Propionib
168	42	43.8	779	2	AAR99843	Thermococ	241	40	41.7	82	4	ABG05138	Novel hum
169	42	43.8	779	8	ADM13795	Thermococ	242	40	41.7	103	6	ABU29579	Protein e
170	42	43.8	779	8	ADM13794	Thermococ	243	40	41.7	107	4	AAG93586	Human ant

244 40 41.7 107 6 ABO27393 Anti-Rh(D) 317 39 40.6 20 8 ADH15516 Gliadin r  
245 40 41.7 114 4 ABO8465 Novel hum 318 39 40.6 65 5 ABO7419 Human ORF  
246 40 41.7 120 3 AAB41201 Human ORF 319 39 40.6 67 4 AAU60358 Propionib  
247 40 41.7 120 8 ADG22437 Cyanophag 320 39 40.6 70 6 ABM56877 Propionib  
248 40 41.7 126 7 ADC95826 E. faeciu 321 39 40.6 70 4 AAG73590 Human col  
249 40 41.7 127 4 AAM24068 Human EST 322 39 40.6 71 5 ABO9115 Human ORF  
250 40 41.7 154 7 ABO61959 Novel hum 323 39 40.6 75 4 AAM23864 Rat EST  
251 40 41.7 165 9 AEA19957 Novel hum 324 39 40.6 88 4 ABB15001 Human ner  
252 40 41.7 169 4 AAU58609 Propionib 325 39 40.6 107 9 AEA40474 Anti-VEGF  
253 40 41.7 169 6 ABM55128 Propionib 326 39 40.6 108 9 ADZ70853 Human Ig  
254 40 41.7 171 9 AEA19958 Human hum 327 39 40.6 108 9 AEB28925 Human ant  
255 40 41.7 180 3 AAB32785 Eucalyptu 328 39 40.6 125 3 AAY70247 C-termina  
256 40 41.7 186 7 ABO68216 Pseudomon 329 39 40.6 135 4 AAM00785 Human bon  
257 40 41.7 191 4 AAU30717 Novel hum 330 39 40.6 137 8 ABO4881 Sea squir  
258 40 41.7 240 2 AAR97195 E.coli ni 331 39 40.6 138 7 ABO74937 Pseudom  
259 40 41.7 240 4 AAU29379 Novel mar 332 39 40.6 143 8 ADX66493 Plant ful  
260 40 41.7 240 6 ABP57430 NADH oxid 333 39 40.6 145 5 ABP63663 Human ORF  
261 40 41.7 263 7 ADB63863 Human pro 334 39 40.6 150 4 ABG08615 Novel hum  
262 40 41.7 285 8 ADX77178 Plant ful 335 39 40.6 161 7 ADM04724 Human pro  
263 40 41.7 292 8 ADX79229 Plant ful 336 39 40.6 207 6 ABU22076 Protein e  
264 40 41.7 296 8 ADY06229 Plant ful 337 39 40.6 215 8 ADH10328 Anti-HIV-  
265 40 41.7 301 3 AAG43485 Arabidops 338 39 40.6 217 8 ADR08459 Human pro  
266 40 41.7 312 8 ADN19945 Bacterial 339 39 40.6 227 8 ADT57705 Plant pol  
267 40 41.7 323 5 ABP56670 Bifidobac 340 39 40.6 240 5 ADH32821 Yeset smo  
268 40 41.7 340 9 ADW46651 Salmonid 341 39 40.6 242 3 AAG28834 Arabidops  
269 40 41.7 345 4 ABG26135 Novel hum 342 39 40.6 242 3 AAG53588 Arabidops  
270 40 41.7 346 6 ADA84007 Human POM 343 39 40.6 244 3 AAG59069 Arabidops  
271 40 41.7 353 3 AAG43484 Arabidops 344 39 40.6 244 8 ADY09152 Plant ful  
272 40 41.7 353 5 ABB92033 Herbicida 345 39 40.6 284 8 ADX87513 Plant ful  
273 40 41.7 380 3 AAG43483 Arabidops 346 39 40.6 298 8 ADT59544 Plant pol  
274 40 41.7 399 3 AAB43375 Human ORF 347 39 40.6 305 8 ADH88967 Escherich  
275 40 41.7 418 7 ABO72078 Pseudomon 348 39 40.6 306 7 ABO81573 Pseudomon  
276 40 41.7 446 8 ADO43442 Lutomyia 349 39 40.6 315 7 ADD27239 Human adi  
277 40 41.7 462 9 AEA20027 Novel hum 350 39 40.6 316 7 ABO73772 Pseudomon  
278 40 41.7 471 9 ABM95232 M. xanthu 351 39 40.6 318 5 ABG60233 Human MAD  
279 40 41.7 504 7 ABM87730 Rice abio 352 39 40.6 318 7 ABG60233 Human MAD  
280 40 41.7 504 8 ADI45243 Rice isop 353 39 40.6 319 2 AAR32193 rfa-2 pro  
281 40 41.7 507 4 ABG20563 Novel hum 354 39 40.6 319 6 ABU14779 Protein e  
282 40 41.7 515 3 AAG46671 Arabidops 355 39 40.6 338 4 ABG99884 S. cinna  
283 40 41.7 596 3 AAG46670 Arabidops 356 39 40.6 341 4 AAB73226 Human pho  
284 40 41.7 614 9 ABM95173 Arabidops 357 39 40.6 357 5 ABP61046 Lactobaci  
285 40 41.7 631 3 AAG46669 Arabidops 358 39 40.6 357 7 ADE12762 L. rhamno  
286 40 41.7 701 8 ADT2767 Bacterial 359 39 40.6 374 9 ADY85524 Catalytic  
287 40 41.7 722 4 AAB92941 Human pro 360 39 40.6 378 4 ABG15295 Novel hum  
288 40 41.7 747 2 AAW89585 Human ATP 361 39 40.6 384 8 ADY23029 Plant ful  
289 40 41.7 747 3 AAB13355 Human ATP 362 39 40.6 384 9 ABM97080 M. xanthu  
290 40 41.7 747 8 ADT89439 Human wil 363 39 40.6 400 8 ADX74229 Plant ful  
291 40 41.7 747 8 ADT89441 Human I39 364 39 40.6 411 8 ADX74229 Plant ful  
292 40 41.7 747 9 ADY58462 Human ATP 365 39 40.6 412 7 ABO74206 Pseudomon  
293 40 41.7 747 9 ADY58464 Human ATP 366 39 40.6 415 8 ADY24642 Plant ful  
294 40 41.7 752 2 AAY14068 Human ABC 367 39 40.6 418 9 AEA27923 Cell surf  
295 40 41.7 752 5 ABP52103 Homo sapi 368 39 40.6 429 3 AAG53587 Arabidops  
296 40 41.7 752 9 ADY70299 Human bet 369 39 40.6 429 3 AAG28833 Arabidops  
297 40 41.7 776 8 ADN75160 JDF-3 pol 370 39 40.6 433 3 AAG28832 Arabidops  
298 40 41.7 776 8 ADN75159 JDF-3 pol 371 39 40.6 433 3 AAG53586 Arabidops  
299 40 41.7 815 8 ADY51572 Arabidops 372 39 40.6 437 8 ADU83348 Human rec  
300 40 41.7 831 8 ADN25201 Bacterial 373 39 40.6 440 6 ABU54546 Human NOV  
301 40 41.7 900 4 ABO26134 Novel hum 374 39 40.6 441 6 ABJ37879 NOX prot  
302 40 41.7 1192 4 ABG23218 Novel hum 375 39 40.6 442 4 AAE11774 Human kin  
303 40 41.7 1305 6 ABU48817 Protein e 376 39 40.6 455 5 AAE25088 Human kin  
304 40 41.7 1770 8 ADN97110 Mouse gua 377 39 40.6 455 5 AAE25088 Human NOV  
305 39.5 41.1 228 6 ABP76813 N. gonorr 378 39 40.6 456 7 ABO69219 Pseudomon  
306 39.5 41.1 244 7 ABM85932 Rice abio 379 39 40.6 466 7 ABO69219 Pseudomon  
307 39.5 41.1 255 6 ABP79809 N. gonorr 380 39 40.6 477 5 ABB79940 Human mem  
308 39.5 41.1 255 6 ABP80516 N. gonorr 381 39 40.6 513 7 ADM03832 Human pro  
309 39.5 41.1 299 8 ADS26731 Bacterial 382 39 40.6 534 6 ABU21476 Protein e  
310 39.5 41.1 299 8 ADS27115 Bacterial 383 39 40.6 537 2 AAY04661 L.lactis  
311 39.5 41.1 303 8 ADS26363 Bacterial 384 39 40.6 546 9 ABM96885 M. xanthu  
312 39.5 41.1 312 9 ADZ75728 Xanthomon 385 39 40.6 548 6 AAE34867 Human kin  
313 39.5 41.1 387 8 ADS30863 Bacterial 386 39 40.6 562 8 ABM83370 Human dia  
314 39 40.6 16 8 ADF91470 Alpha2-gl 387 39 40.6 569 8 ABM83369 Human dia  
315 39 40.6 16 8 ADF91344 High affi 388 39 40.6 575 5 ABB93143 Herbicida  
316 39 40.6 20 8 ADH15852 Gliadin r 389 39 40.6 589 8 ABM83368 Human dia

Adh15516 Gliadin r  
Abp07419 Human ORF  
Aau60358 Propionib  
Abm56877 Propionib  
Aag73590 Human col  
Abp09115 Human ORF  
Abm23864 Rat EST  
Abb15001 Human ner  
Aea40474 Anti-VEGF  
Adz70853 Human Ig  
Aeb28925 Human ant  
Aay70247 C-termina  
Aam00785 Human bon  
Abo4881 Sea squir  
Abo74937 Pseudom  
Adx66493 Plant ful  
Abp63663 Human ORF  
Abg08615 Novel hum  
Adm04724 Human pro  
Abu22076 Protein e  
Adh10328 Anti-HIV-  
Adr08459 Human pro  
Adt57705 Plant pol  
Adh32821 Yeset smo  
Asg28834 Arabidops  
Asg53588 Arabidops  
Asg59069 Arabidops  
Ady09152 Plant ful  
Adx87513 Plant ful  
Adt59544 Plant pol  
Adh88967 Escherich  
Abo81573 Pseudomon  
Add27239 Human adi  
Abo73772 Pseudomon  
Abg60233 Human MAD  
Aam0189 Human NOV  
Aar32193 rfa-2 pro  
Abu14779 Protein e  
Abg99884 S. cinna  
Aab73226 Human pho  
Abp61046 Lactobaci  
Ade12762 L. rhamno  
Ady85524 Catalytic  
Abg15295 Novel hum  
Ady23029 Plant ful  
Abm97080 M. xanthu  
Adx74229 Plant ful  
Abo74206 Pseudomon  
Ady24642 Plant ful  
Aeg75378 Murine IN  
Aea27923 Cell surf  
Aag53587 Arabidops  
Asg28833 Arabidops  
Asg28832 Arabidops  
Asg53586 Arabidops  
Adu83348 Human rec  
Abu54546 Human NOV  
Aej37879 NOX prot  
Aae11774 Human kin  
Aae25088 Novel hum  
Aae25088 Human kin  
Abu54547 Human NOV  
Abo69219 Pseudomon  
Abb79940 Human mem  
Adm03832 Human pro  
Abu21476 Protein e  
Aay04661 L.lactis  
Aay96885 M. xanthu  
Aae34867 Human kin  
Abm83370 Human dia  
Abm83369 Human dia  
Abb93143 Herbicida  
Abm83368 Human dia

390	39	40.6	596	4	ABBY1755
391	39	40.6	602	8	ADQ67657
392	39	40.6	606	4	ABB11013
393	39	40.6	611	4	ABBE0869
394	39	40.6	618	6	ABU54548
395	39	40.6	630	7	ADM05091
396	39	40.6	636	4	ABG27632
397	39	40.6	637	5	ABG66918
398	39	40.6	646	6	ABU28167
399	39	40.6	655	4	AAB20331
400	39	40.6	659	5	Aae22729
401	39	40.6	659	5	ABP51654
402	39	40.6	659	8	ADO44505
403	39	40.6	663	4	ADY37109
404	39	40.6	663	3	ABG29203
405	39	40.6	667	6	ABU14951
406	39	40.6	670	4	RAM24059
407	39	40.6	670	4	AAE03833
408	39	40.6	670	5	ABG64555
409	39	40.6	670	8	ADL77822
410	39	40.6	671	4	ABG17780
411	39	40.6	675	7	ADM04641
412	39	40.6	687	7	ABO67420
413	39	40.6	687	8	ADN17715
414	39	40.6	692	8	ADN24059
415	39	40.6	731	4	AAB79423
416	39	40.6	731	4	AAG91100
417	39	40.6	733	8	ADU66593
418	39	40.6	739	8	ADX91812
419	39	40.6	773	8	ADN75165
420	39	40.6	774	2	RAW29323
421	39	40.6	774	4	AAG78941
422	39	40.6	774	8	ADN75153
423	39	40.6	774	8	ADN75154
424	39	40.6	776	8	ADN75158
425	39	40.6	779	5	ABO07845
426	39	40.6	797	4	ABG29533
427	39	40.6	805	3	AAE70245
428	39	40.6	805	8	AAE25587
429	39	40.6	805	8	ADJ45529
430	39	40.6	807	4	RAM00898
431	39	40.6	915	2	AAE08991
432	39	40.6	915	4	AAU07605
433	39	40.6	915	7	ABR55716
434	39	40.6	915	8	ADJ64379
435	39	40.6	915	9	ADJ07925
436	39	40.6	915	9	ADY91631
437	39	40.6	917	3	AAE58637
438	39	40.6	939	8	ADJ64377
439	39	40.6	950	4	AAO12744
440	39	40.6	958	4	AAO13523
441	39	40.6	958	7	ADC32702
442	39	40.6	959	4	ABE65368
443	39	40.6	984	8	ADN25704
444	39	40.6	1028	7	ADE64073
445	39	40.6	1036	9	ADZ51736
446	39	40.6	1089	5	ABR97258
447	39	40.6	1254	4	ADG07060
448	39	40.6	1254	8	ADG43701
449	39	40.6	1647	4	ABG10750
450	39	40.6	1672	8	ABM81100
451	39	40.6	1690	8	ADG20495
452	39	40.6	1720	5	ADG25089
453	39	40.6	1734	6	ABR42259
454	39	40.6	1734	7	ABM78986
455	39	40.6	1734	8	ADM20497
456	39	40.6	1734	8	ADG20499
457	39	40.6	1734	8	ADG15178
458	39	40.6	1734	8	ADG15178
459	39	40.6	1734	8	ADG15178
460	39	40.6	1734	8	ADG15178
461	39	40.6	1734	8	ADG15178
462	39	40.6	1734	8	ADG15178
463	39	40.6	1734	8	ADG15178
464	39	40.6	1734	8	ADG15178
465	39	40.6	1734	8	ADG15178
466	39	40.6	1734	8	ADG15178
467	39	40.6	1734	8	ADG15178
468	39	40.6	1734	8	ADG15178
469	39	40.6	1734	8	ADG15178
470	39	40.6	1734	8	ADG15178
471	39	40.6	1734	8	ADG15178
472	39	40.6	1734	8	ADG15178
473	39	40.6	1734	8	ADG15178
474	39	40.6	1734	8	ADG15178
475	39	40.6	1734	8	ADG15178
476	39	40.6	1734	8	ADG15178
477	39	40.6	1734	8	ADG15178
478	39	40.6	1734	8	ADG15178
479	39	40.6	1734	8	ADG15178
480	39	40.6	1734	8	ADG15178
481	39	40.6	1734	8	ADG15178
482	39	40.6	1734	8	ADG15178
483	39	40.6	1734	8	ADG15178
484	39	40.6	1734	8	ADG15178
485	39	40.6	1734	8	ADG15178
486	39	40.6	1734	8	ADG15178
487	39	40.6	1734	8	ADG15178
488	39	40.6	1734	8	ADG15178
489	39	40.6	1734	8	ADG15178
490	39	40.6	1734	8	ADG15178
491	39	40.6	1734	8	ADG15178
492	39	40.6	1734	8	ADG15178
493	39	40.6	1734	8	ADG15178
494	39	40.6	1734	8	ADG15178
495	39	40.6	1734	8	ADG15178
496	39	40.6	1734	8	ADG15178
497	39	40.6	1734	8	ADG15178
498	39	40.6	1734	8	ADG15178
499	39	40.6	1734	8	ADG15178
500	39	40.6	1734	8	ADG15178

1798	5	AAW49177	Human	MAA
1798	5	AAE16277	Human kin	
1798	5	ADJ96613	Human mic	
1805	8	ADX91705	Plant ful	
1829	2	AAW29322	DNA polym	
1829	8	ADN59952	Thermococ	
1829	9	ADY81021	Thest DNA	
67	8	ABO55595	Human gen	
128	8	ADQ66298	Novel hum	
160	7	ADC32739	Human nov	
326	8	ADQ79896	Plant pol	
367	8	ADT57468	Plant pol	
390	6	ABR82123	Bacterial	
425	4	ABG12482	Novel hum	
615	7	ABO74457	Pseudomon	
922	8	ADU47668	Streptomy	
1299	7	ADM25903	Hypotherm	
1330	6	ABU26189	Protein e	
1485	2	AAW17887	Photorhab	
1485	2	AAW56546	Toxin tca	
1485	6	ADT21578	Photorhab	
1485	6	ABM70526	Photorhab	
1653	4	AAQ98354	Escherich	
1667	7	ABO65107	Klebsiell	
17	7	ADU14578	Gluten-de	
20	6	AAE34097	Gamma-gli	
20	8	ADH15857	Gliadin r	
20	8	ADH15858	Gliadin r	
20	8	ADH15521	Gliadin r	
20	8	ADH15522	Gliadin r	
20	8	ADH14745	Gliadin r	
28	7	ABD84515	MSRV-1B-p	
38	5	ABH19233	Bacillus	
39	9	ADU35159	Signal pe	
39	9	ADZ80732	Amino aci	
40	4	ABU52680	Human tra	
40	4	ABU53375	Human tes	
40	4	ABU52890	Human tra	
40	4	ABU53138	Human tes	
40	4	AAU58365	Propionib	
57	6	ABM54884	Propionib	
67	2	ADG25146	Codon opt	
80	4	AAE85483	Btt cryst	
80	4	AAE85483	Propionib	
88	4	AAU58167	Propionib	
80	6	ABM54686	Propionib	
84	4	AAQ81407	Human AFP	
84	8	ABO55711	Human gen	
89	4	ABG15087	Novel hum	
92	6	ABP80715	N. gonorr	
92	6	ABU37616	Protein e	
92	6	ABU37715	Protein e	
102	4	ABG15085	Listeria	
104	5	ABR47523	Protein e	
104	6	ABU32821	Fragment	
107	3	AAI84598	Light cha	
107	6	ABR54901	Anti-ido	
107	7	ABE62957	IGG kappa	
107	9	ABE19277	Vlg Fab B	
108	2	AAE75619	Profilin-n	
110	9	ADY96971	Human imm	
112	4	AAU42022	Propionib	
112	4	AAU42022	Propionib	
112	6	ABM38541	Plasmid p	
118	7	ADD13796	Human ORF	
127	5	ABE04113	Profilin-n	
131	2	AAW11487	Zea mays	
131	3	AAQ54494	Corn ZmPR	
131	6	ABE31523	Corn ZmPR	
131	6	AAE31523	Corn ZmPR	
132	3	AAQ50292	Human sec	
136	3	ABM56520	Human pro	
138	9	ADV35165	Signal pe	
139	4	ABH7661	Bovine mam	
141	8	ADY05592	Human mam	

536	38	39.6	141	8	ADX96656	Adx96656 Plant ful	609	38	39.6	245	9	ABM96709	Abm96709 M. xanthu
537	38	39.6	141	8	ADY8763	Ady8763 Plant ful	610	38	39.6	246	3	AAG60406	Aag60406 Arabidops
538	38	39.6	141	8	ADY05934	Ady05934 Plant ful	611	38	39.6	246	3	AAG21409	Aag21409 Arabidops
539	38	39.6	141	8	ADY11436	Ady11436 Plant ful	612	38	39.6	246	3	AAU64120	Aau64120 Propionib
540	38	39.6	141	8	ADY07694	Ady07694 Plant ful	613	38	39.6	253	4	ABM60639	Abm60639 Propionib
541	38	39.6	141	9	ADY35164	Ady35164 Signal pe	614	38	39.6	253	6	ABM60639	Abm60639 Propionib
542	38	39.6	141	9	ADY07222	Ady07222 Amino aci	615	38	39.6	262	4	ABM59592	Abm59592 Human car
543	38	39.6	142	9	ADY25641	Ady25641 Partial C	616	38	39.6	273	8	ABM82948	Abm82948 Human dia
544	38	39.6	143	9	ADY35163	Ady35163 Signal pe	617	38	39.6	274	8	ABM82947	Abm82947 Human dia
545	38	39.6	143	9	ADY35199	Ady35199 Signal pe	618	38	39.6	276	8	ABM82946	Abm82946 Human dia
546	38	39.6	145	9	ADY35162	Ady35162 Signal pe	619	38	39.6	278	8	ABM82945	Abm82945 Human dia
547	38	39.6	146	8	ADY87957	Ady87957 Plant ful	620	38	39.6	279	7	ADB65228	Adb65228 Human pro
548	38	39.6	147	4	AAU67282	Aau67282 Propionib	621	38	39.6	279	8	ABM82944	Abm82944 Human dia
549	38	39.6	147	6	ABM63801	Abm63801 Propionib	622	38	39.6	281	8	ABM82943	Abm82943 Human dia
550	38	39.6	147	8	ADY6721	Ady6721 Plant ful	623	38	39.6	297	7	ABM88726	Abm88726 Rice abio
551	38	39.6	148	4	AAO00329	Aao00329 Human pol	624	38	39.6	298	4	AAU33646	Aau33646 Pseudom
552	38	39.6	148	8	ADY77650	Ady77650 Plant ful	625	38	39.6	298	6	ABU15612	Abu15612 Protein e
553	38	39.6	150	8	ADY07507	Ady07507 Plant ful	626	38	39.6	300	2	ABM99959	Abm99959 B.t.t. co
554	38	39.6	152	9	ADY35161	Ady35161 Signal pe	627	38	39.6	300	7	ABO78549	AbO78549 Pseudom
555	38	39.6	156	8	ADY97226	Ady97226 Plant ful	628	38	39.6	305	8	ABM82940	Abm82940 Human dia
556	38	39.6	156	8	ADY06687	Ady06687 Plant ful	629	38	39.6	308	7	ADJ95128	Adj95128 Novel NOV
557	38	39.6	160	8	ADY71048	Ady71048 Plant ful	630	38	39.6	308	8	ADN04269	Adn04269 Antipsoi
558	38	39.6	173	8	ADT90716	Adt90716 Human pro	631	38	39.6	310	8	ABM82942	Abm82942 Human dia
559	38	39.6	177	8	ADY04229	Ady04229 Plant ful	632	38	39.6	312	3	AAG35609	Aag35609 Arabidops
560	38	39.6	179	9	ADY35160	Ady35160 Signal pe	633	38	39.6	313	3	ABM82941	Abm82941 Human dia
561	38	39.6	181	3	ADY92231	Ady92231 Plant ful	634	38	39.6	322	7	AAG35608	Aag35608 Arabidops
562	38	39.6	182	3	AAG54493	Aag54493 Zea mays	635	38	39.6	322	7	ADF58900	Adf58900 Human pol
563	38	39.6	198	6	ABU31293	Abu31293 Protein e	636	38	39.6	322	8	ADZ14522	Adz14522 Human SVC
564	38	39.6	205	8	ADL90209	Adl90209 Human enz	637	38	39.6	326	8	ADS23755	Ads23755 Bacterial
565	38	39.6	206	9	AEA39536	Aea39536 Nematode	638	38	39.6	329	1	AAP94618	Aap94618 Sequence
566	38	39.6	207	7	ABO73028	AbO73028 Pseudom	639	38	39.6	329	4	AAE12008	Aae12008 Streptomy
567	38	39.6	220	3	AAH84583	Aah84583 Amino aci	640	38	39.6	329	4	AAE12004	Aae12004 Streptomy
568	38	39.6	220	3	AAH44776	Aah44776 Short spl	641	38	39.6	329	6	ABG73757	Abg73757 S. clavul
569	38	39.6	220	3	AAH68710	Aah68710 A human p	642	38	39.6	329	6	ABG73761	Abg73761 S. clavul
570	38	39.6	220	4	ABM50978	Abm50978 Human PRO	643	38	39.6	332	7	ABO73894	AbO73894 Pseudom
571	38	39.6	220	5	AAU86158	Aau86158 Human PRO	644	38	39.6	335	8	ADY78437	Ady78437 Plant ful
572	38	39.6	220	5	ABM84975	Abm84975 Human PRO	645	38	39.6	341	8	ADY08488	Ady08488 Plant ful
573	38	39.6	220	5	ABG30698	Abg30698 Human art	646	38	39.6	343	5	ABM78776	Abm78776 Soybean f
574	38	39.6	220	5	ABM82388	Abm82388 Human neu	647	38	39.6	345	4	AAU40084	Aau40084 Human pol
575	38	39.6	220	5	ABM95581	Abm95581 Human ang	648	38	39.6	347	4	AAU41481	Aau41481 Propionib
576	38	39.6	220	5	ABO22940	AbO22940 Human foe	649	38	39.6	347	6	ABM38000	Abm38000 Propionib
577	38	39.6	220	6	ABU56702	Abu56702 Lung canc	650	38	39.6	354	8	ADT90717	Adt90717 Human pro
578	38	39.6	220	6	ABU56539	Abu56539 Lung canc	651	38	39.6	373	3	AAG25607	Aag25607 Arabidops
579	38	39.6	220	6	ABU56703	Abu56703 Lung canc	652	38	39.6	373	5	ABM92899	Abm92899 Herbicida
580	38	39.6	220	6	ABU56540	Abu56540 Lung canc	653	38	39.6	374	5	ABM48190	Abm48190 Listeria
581	38	39.6	220	6	ABU71444	Abu71444 Human neo	654	38	39.6	374	6	ABU32902	Abu32902 Protein e
582	38	39.6	220	7	ADD10607	Add10607 Human sec	655	38	39.6	382	6	ABJ25564	Abj25564 Aspergill
583	38	39.6	220	7	ADD11567	Add11567 Human sec	656	38	39.6	400	4	AAU42254	Aau42254 Propionib
584	38	39.6	220	7	ADD37360	Add37360 Human sec	657	38	39.6	400	6	ABM38773	Abm38773 Propionib
585	38	39.6	220	8	ADH43751	Adh43751 Human PRO	658	38	39.6	404	8	ADT59641	Adt59641 Plant pol
586	38	39.6	220	8	ADK83096	Adk83096 Human PRO	659	38	39.6	409	4	AAU36106	Aau36106 Klebsiell
587	38	39.6	220	8	ADL16439	Adl16439 Human pre	660	38	39.6	410	4	ABM62935	Abm62935 Drosophil
588	38	39.6	220	8	ADU46758	Adu46758 Human pre	661	38	39.6	410	7	ABO77908	AbO77908 Pseudom
589	38	39.6	220	9	ADV35145	Adv35145 Human pre	662	38	39.6	417	4	AAU78377	Aau78377 Human pro
590	38	39.6	220	9	ADZ80734	Adz80734 Amino aci	663	38	39.6	417	4	AAU79361	Aau79361 Human pro
591	38	39.6	224	6	ABU22613	Abu22613 Protein e	664	38	39.6	427	9	ABM95053	Abm95053 M. xanthu
592	38	39.6	225	3	AAH32119	Aah32119 Maize glu	665	38	39.6	430	3	AAG38567	Aag38567 Arabidops
593	38	39.6	226	7	ABO80943	AbO80943 Pseudom	666	38	39.6	430	3	AAG05065	Aag05065 Arabidops
594	38	39.6	227	8	AAU58149	Aau58149 Plant pol	667	38	39.6	432	3	AAG38566	Aag38566 Arabidops
595	38	39.6	228	4	AAU51500	Aau51500 Propionib	668	38	39.6	432	3	AAG05064	Aag05064 Arabidops
596	38	39.6	228	6	ABM48019	Abm48019 Propionib	669	38	39.6	432	8	AAU74809	Aau74809 Thale cre
597	38	39.6	230	4	AAU30118	Aau30118 Novel hum	670	38	39.6	436	3	AAG05063	Aag05063 Arabidops
598	38	39.6	226	7	ABO80943	AbO80943 Pseudom	671	38	39.6	436	7	ABO74759	AbO74759 Pseudom
599	38	39.6	227	8	AAU58149	Aau58149 Plant pol	672	38	39.6	448	3	AAG38565	Aag38565 Arabidops
600	38	39.6	228	4	AAU51500	Aau51500 Propionib	673	38	39.6	448	7	ABO78751	AbO78751 Pseudom
601	38	39.6	228	6	ABM48019	Abm48019 Propionib	674	38	39.6	449	8	ADN17696	Adn17696 Bacterial
602	38	39.6	230	4	AAU30118	Aau30118 Novel hum	675	38	39.6	451	6	ADY27983	Ady27983 Protein e
603	38	39.6	231	7	ABO81216	AbO81216 Pseudom	676	38	39.6	461	8	ADY13657	Ady13657 Plant ful
604	38	39.6	232	3	AAH50923	Aah50923 Human fet	677	38	39.6	478	2	AAR47586	Aar47586 Human bon
605	38	39.6	232	8	ADL90198	Adl90198 Human enz	678	38	39.6	478	2	AAR47586	Aar47586 Human bon
606	38	39.6	240	7	ABO64794	AbO64794 Klebsiell	679	38	39.6	478	3	AAH92023	Aah92023 Human bon
607	38	39.6	242	3	AAG60407	Aag60407 Arabidops	680	38	39.6	478	3	ADN39338	Adn39338 Cancer/an
608	38	39.6	242	3	AAG21410	Aag21410 Arabidops	681	38	39.6	478	7	ABO82709	AbO82709 Pseudom

682	38	39.6	478	7	AEb86462	Human BMP	755	38	39.6	644	2	AAW34837	Novel Cry
683	38	39.6	478	8	ADH11577	Human bon	756	38	39.6	644	2	AAW34826	Novel Cry
684	38	39.6	478	8	ADQ19527	Human sof	757	38	39.6	644	2	AAW34828	Novel Cry
685	38	39.6	478	9	ADY15678	PRO poly	758	38	39.6	644	2	AAW34840	Novel Cry
686	38	39.6	490	8	ADT90715	Human pro	759	38	39.6	644	2	AAW34822	Novel Cry
687	38	39.6	493	7	ABO61738	Klebsiell	760	38	39.6	644	2	AAW34818	Novel Cry
688	38	39.6	497	4	AAU18149	Novel hum	761	38	39.6	644	2	AAW34820	Novel Cry
689	38	39.6	497	4	AAU17035	Human nov	762	38	39.6	644	2	AAW34833	Novel Cry
690	38	39.6	497	4	ABB10508	Human CDN	763	38	39.6	644	2	AAW34811	Novel Cry
691	38	39.6	497	4	AAU19960	Novel hum	764	38	39.6	644	2	AAW34819	Novel Cry
692	38	39.6	497	4	AAU87636	Novel cen	765	38	39.6	644	2	AAW34824	Novel Cry
693	38	39.6	497	5	ABJ05776	Novel hum	766	38	39.6	644	2	AAW34812	Novel Cry
694	38	39.6	497	5	ABP67095	Human pol	767	38	39.6	644	2	AAW34813	Novel Cry
695	38	39.6	497	8	ADI54951	Novel hum	768	38	39.6	644	2	AAW34817	Novel Cry
696	38	39.6	500	7	ADD13792	Plasmid p	769	38	39.6	644	4	AAW34819	Novel Cry
697	38	39.6	512	5	ABBS5533	Lactococc	770	38	39.6	644	6	ABP72637	Bacillus
698	38	39.6	529	7	ABO83094	Pseudomon	771	38	39.6	644	6	AAO27315	Bacillus
699	38	39.6	531	6	ADA34890	Acinetoba	772	38	39.6	644	2	AAW34784	An artifi
700	38	39.6	553	4	ABB52702	Escherich	773	38	39.6	646	8	ADW76603	Human typ
701	38	39.6	556	4	ABG07378	Novel hum	774	38	39.6	650	4	AAW38706	Human pol
702	38	39.6	577	8	ADG45034	Bacterial	775	38	39.6	651	2	AAW23197	Amino aci
703	38	39.6	578	4	AAW38707	Human pol	776	38	39.6	651	2	AAW23200	Amino aci
704	38	39.6	578	9	ADV77043	Huntingto	777	38	39.6	651	2	AAW23199	Amino aci
705	38	39.6	584	4	AAW41870	Human pol	778	38	39.6	652	2	AAW23214	Amino aci
706	38	39.6	585	7	ABO67351	Klebsiell	779	38	39.6	652	8	ADR89423	cry3Aa1
707	38	39.6	596	6	ABP72639	Modified	780	38	39.6	652	9	ADY59863	Bacillus
708	38	39.6	597	2	AAW00334	Bacillus	781	38	39.6	652	9	AEA37685	Longicorn
709	38	39.6	597	3	AAW69666	Wild-type	782	38	39.6	653	6	ABU32366	Protein e
710	38	39.6	597	6	ABP72638	Cry3A tox	783	38	39.6	669	6	ABU26164	Aspergill
711	38	39.6	597	6	ABP72646	Modified	784	38	39.6	669	6	ABU28010	Protein e
712	38	39.6	597	6	ABP72642	Modified	785	38	39.6	680	8	ADX94704	Plant ful
713	38	39.6	598	6	ABP72640	Modified	786	38	39.6	684	7	ADB64264	Human pro
714	38	39.6	600	6	ABP72644	Modified	787	38	39.6	684	9	ADY18110	PRO poly
715	38	39.6	601	6	ABP72643	Modified	788	38	39.6	685	7	ADE28201	Human MDD
716	38	39.6	602	6	ABP72645	Modified	789	38	39.6	687	7	ABO63351	Klebsiell
717	38	39.6	604	8	ADT90724	Human pro	790	38	39.6	697	7	ADH88031	Enterococ
718	38	39.6	606	6	ABP72641	Modified	791	38	39.6	720	8	ADS23416	Bacterial
719	38	39.6	610	2	AAW05537	Synthetic	792	38	39.6	730	5	AAU74439	Worm prot
720	38	39.6	610	2	AAW85482	Btt synth	793	38	39.6	742	7	ASO61481	LDL recep
721	38	39.6	610	2	AAW00333	Bacillus	794	38	39.6	749	5	ABO4874	LDL recep
722	38	39.6	610	2	ADG25144	Codon opt	795	38	39.6	759	4	AAW40492	Human pol
723	38	39.6	610	3	AAW69667	Bacillus	796	38	39.6	759	4	AAW40493	Human pol
724	38	39.6	613	4	ABB71965	Drosophil	797	38	39.6	760	7	ADN95128	Human LEC
725	38	39.6	613	7	ABO74408	Pseudomon	798	38	39.6	771	5	ABO5675	Human cel
726	38	39.6	620	7	ADN95361	Human BEC	799	38	39.6	771	8	ADU98566	Protein f
727	38	39.6	620	8	ADP54705	Human PRO	800	38	39.6	773	8	ADN75164	Tgo polym
728	38	39.6	644	1	AAW70085	Sequence	801	38	39.6	774	8	ADN75152	KOD polym
729	38	39.6	644	1	AAW82487	Delta end	802	38	39.6	774	8	ADN75152	KOD polym
730	38	39.6	644	1	AAW80467	Bacillus	803	38	39.6	775	7	ADG64508	Family B
731	38	39.6	644	1	AAW94679	Amino aci	804	38	39.6	775	8	ADO58193	Mutant Pf
732	38	39.6	644	1	AAW95585	M-7 cryst	805	38	39.6	775	8	ADO58193	Mutant Pf
733	38	39.6	644	2	AAW39751	Delta end	806	38	39.6	775	8	ADN75168	Pfu polym
734	38	39.6	644	2	AAW56697	CryIIIA i	807	38	39.6	775	8	ADN59954	9on Therm
735	38	39.6	644	2	AAW56698	CryIIB in	808	38	39.6	775	8	ADQ26674	Mutant Pf
736	38	39.6	644	2	AAW99960	B.t.t.co	809	38	39.6	775	8	ADQ26677	Mutant Pf
737	38	39.6	644	2	AAW99958	Bacillus	810	38	39.6	775	8	ADT07152	Pfu mutan
738	38	39.6	644	2	AAW34816	Novel Cry	811	38	39.6	775	8	ADT07149	Pfu mutan
739	38	39.6	644	2	AAW34821	Novel Cry	812	38	39.6	775	9	ADY63793	9degreeen
740	38	39.6	644	2	AAW34829	Novel Cry	813	38	39.6	775	9	ADY63797	9degreeen
741	38	39.6	644	2	AAW34825	Novel Cry	814	38	39.6	775	9	ADY63801	9degreeen
742	38	39.6	644	2	AAW34834	Novel Cry	815	38	39.6	775	9	ADY63795	9degreeen
743	38	39.6	644	2	AAW34838	Novel Cry	816	38	39.6	775	9	ADY63803	9degreeen
744	38	39.6	644	2	AAW34814	Novel Cry	817	38	39.6	775	9	ADY63799	9degreeen
745	38	39.6	644	2	AAW34830	Novel Cry	818	38	39.6	775	9	ADY81028	9degn DNA
746	38	39.6	644	2	AAW34827	Novel Cry	819	38	39.6	776	9	ADY81038	9on-7 DNA
747	38	39.6	644	2	AAW34835	Novel Cry	820	38	39.6	776	8	ADO58206	Mutant JD
748	38	39.6	644	2	AAW34815	Novel Cry	821	38	39.6	776	8	ADN75161	JDF-3 pol
749	38	39.6	644	2	AAW34823	Novel Cry	822	38	39.6	776	8	ADN75156	JDF-3 pol
750	38	39.6	644	2	AAW34836	Novel Cry	823	38	39.6	776	8	ADQ26687	Mutant JD
751	38	39.6	644	2	AAW34839	Novel Cry	824	38	39.6	776	8	ADT07162	Pfu mutan
752	38	39.6	644	2	AAW34841	Novel Cry	825	38	39.6	777	6	ABM70298	Photorhab
753	38	39.6	644	2	AAW34831	Novel Cry	826	38	39.6	797	5	ADK34903	hum
754	38	39.6	644	2	AAW34832	Novel Cry	827	38	39.6	798	3	AAW52699	Aspergill



828	38	39.6	828	6	ABU39423	Abu39423	Protein e	901	38	39.6	3570	8	ADH71310	Adh71310	Human pro
829	38	39.6	828	8	ADN18248	Adn18248	Bacterial	902	38	39.6	3570	8	ADH71308	Adh71308	Human pro
830	38	39.6	829	8	ADS25721	Ads25721	Bacterial	903	38	39.6	3570	8	ADH71300	Adh71300	Human pro
831	38	39.6	834	8	ADS22448	Ads22448	Bacterial	904	38	39.6	3570	8	ADH71306	Adh71306	Human pro
832	38	39.6	836	4	AAB45914	Aab45914	S. enteri	905	38	39.6	3570	8	ADO09972	Ado09972	Human NOV
833	38	39.6	846	8	ADO61707	Ado61707	Transcrip	906	38	39.6	3570	8	ADO09840	Ado09840	Human NOV
834	38	39.6	850	8	ADS26200	Ads26200	Bacterial	907	38	39.6	3571	5	AAE20146	Aae20146	Human C3b
835	38	39.6	853	3	AAB26917	Aab26917	Large sub	908	38	39.6	3571	8	ADL22512	Adl22512	Human CNG
836	38	39.6	853	7	ABU62618	Abu62618	Pseudomon	909	38	39.6	3571	8	ADS97984	Ads97984	Protein f
837	38	39.6	914	2	AAR15785	Aar15785	B. thuring	910	38	39.6	4019	4	AAE13839	Aae13839	Human lun
838	38	39.6	917	8	ADK52123	Adk52123	Human ato	911	38	39.6	4019	7	ADDE6733	Adde6733	Human lun
839	38	39.6	917	8	ADQ80402	Adq80402	KIAA1543	912	38	39.6	4019	7	ADDE87987	Adde87987	Human lun
840	38	39.6	934	6	ABU40783	Abu40783	Protein e	913	38	39.6	4025	5	ABP69736	Abp69736	Human pol
841	38	39.6	935	4	ABG10172	Abg10172	Novel hum	914	38	39.6	4025	5	ADU18074	Adu18074	Human can
842	38	39.6	937	7	ADF05228	Adf05228	Bacterial	915	37.5	39.1	58	5	ABP05637	Abp05637	Human ORF
843	38	39.6	956	2	AAR15784	Aar15784	B. thuring	916	37.5	39.1	60	4	AAU41788	Aau41788	Propionib
844	38	39.6	967	5	ADR41532	Adr41532	Human CD-	917	37.5	39.1	60	6	ABM38307	Abm38307	Propionib
845	38	39.6	976	8	ADS27751	Ads27751	Bacterial	918	37.5	39.1	61	5	ABP06432	Abp06432	Human ORF
846	38	39.6	979	4	AGG66398	Agg66398	Receptor	919	37.5	39.1	76	4	AAU40101	Aau40101	Propionib
847	38	39.6	1000	8	ADK98523	Adk98523	S. Bulboca	920	37.5	39.1	76	6	ABM36620	Abm36620	Propionib
848	38	39.6	1028	7	ADE64071	Ade64071	Rat Prote	921	37.5	39.1	77	4	AAI19818	Aai19818	Peptide #
849	38	39.6	1031	7	ABM85226	Abm85226	Mouse pro	922	37.5	39.1	77	4	ABM39693	Abm39693	Peptide #
850	38	39.6	1039	7	ABM85227	Abm85227	Human pro	923	37.5	39.1	77	4	ABM33269	Abm33269	Peptide #
851	38	39.6	1044	6	ABU62348	Abu62348	Mouse nuc	924	37.5	39.1	77	4	ABM24353	Abm24353	Protein #
852	38	39.6	1044	8	ADG28516	Adg28516	Bovine nu	925	37.5	39.1	77	4	AAU73047	Aau73047	Human bon
853	38	39.6	1050	4	ABG10171	Abg10171	Novel hum	926	37.5	39.1	77	4	AAAG60411	Aaag60411	Human bra
854	38	39.6	1063	4	AAU39991	Aau39991	Human pol	927	37.5	39.1	77	4	ABG54761	Abg54761	Human liv
855	38	39.6	1066	4	ABP37987	Abp37987	Human GS9	928	37.5	39.1	77	5	ABG42891	Abg42891	Human pep
856	38	39.6	1092	7	ADH88406	Adh88406	Enterococ	929	37.5	39.1	96	5	AAO15965	Aao15965	DNA polym
857	38	39.6	1100	2	AAR15783	Aar15783	B. thuring	930	37.5	39.1	108	6	ABP60623	Abp60623	zinc fing
858	38	39.6	1101	4	AAB82299	Aab82299	Wheat sta	931	37.5	39.1	108	7	ADU42063	Adu42063	Human can
859	38	39.6	1139	4	AAU18126	Aau18126	Novel hum	932	37.5	39.1	156	4	AAU16255	Aau16255	Human nov
860	38	39.6	1139	4	AAU16963	Aau16963	Human cdn	933	37.5	39.1	156	6	AAU55324	Aau55324	Human nov
861	38	39.6	1139	4	ABU10326	Abu10326	Human nov	934	37.5	39.1	169	4	AAU42685	Aau42685	Propionib
862	38	39.6	1139	4	AAU19902	Aau19902	Novel hum	935	37.5	39.1	169	6	ABM39204	Abm39204	Propionib
863	38	39.6	1139	4	AAU87355	Aau87355	Novel cen	936	37.5	39.1	244	4	ABG44138	Abg44138	Novel hum
864	38	39.6	1139	5	ABU05753	Abu05753	Novel hum	937	37.5	39.1	294	5	ABM89627	Abm89627	Human pol
865	38	39.6	1139	5	ABP66913	Abp66913	Human pol	938	37.5	39.1	351	8	ADJ50995	Adj50995	Cancer re
866	38	39.6	1139	8	AD154670	Adi154670	Novel hum	939	37.5	39.1	364	8	ADJ37239	Adj37239	Human nuc
867	38	39.6	1197	7	ABO82087	Abu82087	Pseudomon	940	37.5	39.1	476	7	ADB64886	Adb64886	Human pro
868	38	39.6	1251	5	AAU74829	Aau74829	Human REP	941	37.5	39.1	504	7	ADB64735	Adb64735	Human pro
869	38	39.6	1292	5	ABB77986	Abb77986	Amino aci	942	37.5	39.1	510	4	AAU41862	Aau41862	Human pol
870	38	39.6	1316	4	AAB27248	Aab27248	Maize RAD	943	37.5	39.1	525	6	ABJ26675	Abj26675	Seed deve
871	38	39.6	1376	4	ABB59729	Abb59729	Drosophil	944	37.5	39.1	558	4	AAAG5577	Aaag5577	Human hmi
872	38	39.6	1422	4	ABG02345	Abg02345	Novel hum	945	37.5	39.1	558	4	AAU40076	Aau40076	Human pol
873	38	39.6	1487	5	AAO14351	Aao14351	Equine he	946	37.5	39.1	558	4	AAU92849	Aau92849	Human pro
874	38	39.6	1487	5	AAO14354	Aao14354	Equine he	947	37.5	39.1	558	4	AAU93732	Aau93732	Human pro
875	38	39.6	1487	5	AAO14353	Aao14353	Equine he	948	37.5	39.1	558	5	AAO18891	Aao18891	Human ova
876	38	39.6	1487	5	AAO14356	Aao14356	Equine he	949	37.5	39.1	620	8	ADN99715	Adn99715	Novel hum
877	38	39.6	1487	5	AAO14352	Aao14352	Equine he	950	37.5	39.1	620	8	ADN99715	Adn99715	Novel hum
878	38	39.6	1487	5	AAO14355	Aao14355	Equine he	951	37.5	39.1	625	8	ADI45445	Adi45445	Rice isop
879	38	39.6	1487	5	AAU93954	Aau93954	Human pol	952	37.5	39.1	659	8	ADN99714	Adn99714	Novel hum
880	38	39.6	1497	8	ADL32120	Adl32120	Human pro	953	37.5	39.1	684	8	ADN20726	Adn20726	Bacterial
881	38	39.6	1498	7	ADC86869	Adc86869	Human GPC	954	37.5	39.1	709	4	AAU93837	Aau93837	Human pol
882	38	39.6	1568	8	ADO00979	Ado00979	Mouse hom	955	37.5	39.1	709	8	ADL31876	Adl31876	Human pro
883	38	39.6	1964	2	AAW95557	Aaw95557	Mus muscu	956	37.5	39.1	770	8	ADN05763	Adn05763	Mouse Not
884	38	39.6	1964	8	ADP74016	Adp74016	Murine No	957	37.5	39.1	855	7	ADC08028	Adc08028	Rice prot
885	38	39.6	1964	8	ADT49888	Adt49888	Murine NO	958	37.5	39.1	886	8	ADR05762	Adr05762	Human RNA
886	38	39.6	2050	5	AAAG68264	Aag68264	Human POL	959	37.5	39.1	887	2	AAU28995	Aau28995	Tumour su
887	38	39.6	2050	6	ABU12097	Abu12097	Novel hum	960	37.5	39.1	887	2	AAU15344	Aau15344	Tumour su
888	38	39.6	2050	7	AAE38812	Aae38812	Human com	961	37.5	39.1	887	5	AAO21729	Aao21729	Human AD7
889	38	39.6	2050	8	ADH71296	Adh71296	Human pro	962	37.5	39.1	887	9	ADL82973	Adl82973	Human PRO
890	38	39.6	2706	8	ADS97987	Ads97987	Protein f	963	37	38.5	20	8	ADH15854	Adh15854	GLIadin r
891	38	39.6	3557	7	ADN39112	Adn39112	Cancer/an	964	37	38.5	20	8	ADH15518	Adh15518	GLIadin r
892	38	39.6	3557	7	ADN39979	Adn39979	Cancer/an	965	37	38.5	56	5	ABG30476	Abg30476	Human ant
893	38	39.6	3565	8	ADH71282	Adh71282	Human pro	966	37	38.5	74	8	ADX94926	Adx94926	Plant ful
894	38	39.6	3568	5	ABJ10590	Abj10590	Human nov	967	37	38.5	75	8	ADX71090	Adx71090	Plant ful
895	38	39.6	3568	8	ADH71298	Adh71298	Human pro	968	37	38.5	75	5	ABP31093	Abp31093	Human ORF
896	38	39.6	3568	8	ADO09842	Ado09842	Human NOV	969	37	38.5	77	9	AEA21086	Aea21086	Novel hum
897	38	39.6	3570	5	ABJ10589	Abj10589	Human nov	970	37	38.5	80	4	AAU42555	Aau42555	Propionib
898	38	39.6	3570	7	AAE37944	Aae37944	Human CGD	971	37	38.5	80	6	AAU39074	Adm39074	Propionib
899	38	39.6	3570	8	ADH71302	Adh71302	Human pro	972	37	38.5	82	5	ABP64482	Abp64482	Human ORF
900	38	39.6	3570	8	ADH71304	Adh71304	Human pro	973	37	38.5	86	4	AAU61523	Aau61523	Propionib

974 37 38.5 86 6 ABM58042  
975 37 38.5 89 6 ABM64719  
976 37 38.5 90 4 ABG27586  
977 37 38.5 93 8 ADX94899  
978 37 38.5 94 5 ABP07464  
979 37 38.5 96 8 ADH71010  
980 37 38.5 100 8 ADX94251  
981 37 38.5 104 4 ABG15093  
982 37 38.5 105 2 AAW52239  
983 37 38.5 107 5 ABG30494  
984 37 38.5 107 8 ABM79492  
985 37 38.5 107 8 ADP22110  
986 37 38.5 107 8 ADQ77010  
987 37 38.5 107 9 AEA40482  
988 37 38.5 108 8 ADS16543  
989 37 38.5 109 7 ADL91335  
990 37 38.5 109 8 ADO36362  
991 37 38.5 109 8 ADO75235  
992 37 38.5 122 4 ABG08901  
993 37 38.5 123 4 AAO00225  
994 37 38.5 128 4 AAU63650  
995 37 38.5 128 6 ABM60169  
996 37 38.5 129 3 AAB41339  
997 37 38.5 130 4 ABG17269  
998 37 38.5 134 7 ABM85558  
999 37 38.5 135 4 AAU54466  
1000 37 38.5 135 6 ABM50985

## ALIGNMENTS

RESULT 1  
ID ADD26060 standard; peptide; 17 AA.  
XX  
AC ADD26060;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE CRX peptide used in the invention.  
XX  
KW CRYPTO; Cytostatic; tumour; cell proliferation; antibody.  
XX  
OS Synthetic.  
XX  
PN WO2003083041-A2.  
XX  
PD 09-OCT-2003.  
XX  
PF 01-OCT-2002; 2002WO-US031462.  
XX  
PR 22-MAR-2002; 2002US-0367002P.  
PR 24-APR-2002; 2002WO-US011950.  
XX  
XX (SANI/) SANICOLA-NADEL M.  
PA (ADKI/) ADKINS H.  
PA (MIKL/) MIKLASZ S D.  
PA (RAYH/) RAYHORN P.  
PA (SCH/) SCHIFFER S G.  
PA (WILL/) WILLIAMS K.  
XX  
XX Sanicola-Nadel M, Adkins H, Miklasz SD, Rayhorn P, Schiffer SG;  
PI Williams K;  
XX  
XX WPI; 2003-779452/73.  
XX  
XX New anti-Cripto antibodies, useful in the therapy of malignant or benign  
PT tumors of mammals, or for treating a mammal for a condition associated  
PT with undesired cell proliferation.  
XX  
PS Example 2; SEQ ID NO 3; 89pp; English.  
XX

CC The present invention relates to an antibody that specifically binds to  
CC Cripto, to an epitope in the ligand/receptor binding domain of Cripto, or  
CC to an epitope comprising the domain spanning amino acid residues 46-62.  
CC The antibodies or compositions comprising them are useful for inhibiting  
CC growth of tumour cells, treating a mammal having a tumor that  
CC overexpresses Cripto, or treating a mammal for a condition associated  
CC with undesired cell proliferation. The antibodies are useful in the  
CC therapy of malignant or benign tumors of mammals where growth of the  
CC tumor is at least partially dependent upon Cripto. The present sequence  
CC represents a CRX peptide of the invention.  
XX  
SQ Sequence 17 AA;  
Query Match 100.0%; Score 96; DB 7; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.3e-07; Mismatches 0; Gaps 0;  
Matches 17; Conservative 0;  
QY 1 FRDSDIWPQEPAIRPR 17  
DB 1 FRDSDIWPQEPAIRPR 17  
|||||  
RESULT 2  
ABB77111  
ID ABB77111 standard; protein; 129 AA.  
XX  
AC ABB77111;  
XX  
DT 08-OCT-2002 (first entry)  
XX  
DE Human Cripto-1 C-terminal truncated form mutant CR38 (cC)T88A.  
XX  
KW Human; Cripto-1; CR-1; mutant; tumour; cell proliferation.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 48  
FT /note= "Wild type Thr substituted by Ala"  
XX  
XX WO200222808-A2.  
XX  
PD 21-MAR-2002.  
XX  
PF 18-SEP-2001; 2001WO-US029066.  
XX  
PR 18-SEP-2000; 2000US-0233148P.  
XX  
XX (BIOJ ) BIOGEN INC.  
XX  
XX Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
XX  
XX WPI; 2002-339868/37.  
XX  
XX New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful  
PT for treating cell proliferation, especially cancer, comprises amino acid  
PT change that prevents fucosylation at Thr88.  
XX  
PS Disclosure; Page 36; 41pp; English.  
XX  
XX The sequence represents a C-terminal truncated form of human Cripto-1,  
CC where the threonine residue at position 48 has been substituted with an  
CC alanine residue. The invention relates to a novel mutant CRIPTO  
CC polypeptide, or its functional fragment, having at least one amino acid  
CC alteration at positions 86, 87 or 88. The mutant polypeptide, or its  
CC chimera, is used to inhibit growth of tumour cells, in vivo or in vitro,  
CC particularly for treating breast, ovarian, renal, colorectal, uterine,  
CC prostatic, lung, bladder or central nervous system cancers, melanoma and  
CC leukaemia, also generally for treating undesired cell proliferation  
XX  
SQ Sequence 129 AA;

```
Query Match      100.0%; Score 96; DB 5; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSDIWPQEEPAIRPR 17
Db      6 FRDSDIWPQEEPAIRPR 22

RESULT 3
ABB77124
ID ABB77124 standard; protein; 129 AA.
XX
AC ABB77124;
XX
DT 08-OCT-2002 (first entry)
XX
DE Human Cripto-3 C-terminal truncated form mutant CR-338(cc)T98A.
XX
KW Human; Cripto-1; CR-1; mutant; tumour; Cripto-3; CR-3;
XX cell proliferation.
XX
OS Homo sapiens.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 48
FT /note= "Wild type Thr substituted by Ala"
XX
PN WO200222808-A2.
XX
PD 21-MAR-2002.
XX
PF 18-SEP-2001; 2001WO-US029066.
XX
PR 18-SEP-2000; 2000US-0233148P.
XX
PA (BIOJ ) BIOGEN INC.
XX
PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;
XX WPI; 2002-339868/37.
XX
DR New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful
XX for treating cell proliferation, especially cancer, comprises amino acid
XX change that prevents fucosylation at Thr88.
XX
PS Disclosure; Page 39; 41pp; English.
XX
CC The sequence represents a C-terminal truncated form of human Cripto-3,
XX where the threonine residue at position 48 has been substituted with an
XX alanine residue. The invention relates to a novel mutant CRIPTO
XX polypeptide, or its functional fragment, having at least one amino acid
XX alteration at positions 86, 87 or 88. The mutant polypeptide, or its
XX chimera, is used to inhibit growth of tumour cells, in vivo or in vitro,
XX particularly for treating breast, ovarian, renal, colorectal, uterine,
XX prostatic, lung, bladder or central nervous system cancers, melanoma and
XX leukaemia, also generally for treating undesired cell proliferation
XX
SQ Sequence 129 AA;

Query Match      100.0%; Score 96; DB 5; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSDIWPQEEPAIRPR 17
Db      6 FRDSDIWPQEEPAIRPR 22

RESULT 4
ABB77104
ID ABB77104 standard; protein; 129 AA.
XX
AC ABB77104;
XX
DT 08-OCT-2002 (first entry)
XX
DE Human Cripto-1 C-terminal truncated form (CR38(cc)).
XX
KW Human; Cripto-1; CR-1; mutant; tumour; cell proliferation.
XX
OS Homo sapiens.
XX
PN WO200222808-A2.
XX
PD 21-MAR-2002.
XX
PF 18-SEP-2001; 2001WO-US029066.
XX
PR 18-SEP-2000; 2000US-0233148P.
XX
PA (BIOJ ) BIOGEN INC.
XX
PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;
XX WPI; 2002-339868/37.
XX
DR New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful
XX for treating cell proliferation, especially cancer, comprises amino acid
XX change that prevents fucosylation at Thr88.
XX
PS Claim 2; Page 35; 41pp; English.
XX
CC The sequence represents a C-terminal truncated form of human Cripto-1.
XX The invention relates to a novel mutant CRIPTO polypeptide, or its
XX functional fragment, having at least one amino acid alteration at
XX positions 86, 87 or 88. The mutant polypeptide, or its chimera, is used
XX to inhibit growth of tumour cells, in vivo or in vitro, particularly for
XX treating breast, ovarian, renal, colorectal, uterine, prostatic, lung,
XX bladder or central nervous system cancers, melanoma and leukaemia, also
XX generally for treating undesired cell proliferation
XX
SQ Sequence 129 AA;

Query Match      100.0%; Score 96; DB 5; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRDSDIWPQEEPAIRPR 17
Db      6 FRDSDIWPQEEPAIRPR 22

RESULT 5
ABB77117
ID ABB77117 standard; protein; 129 AA.
XX
AC ABB77117;
XX
DT 08-OCT-2002 (first entry)
XX
DE Human Cripto-3 C-terminal truncated form (CR-338(cc)).
XX
KW Human; Cripto-1; CR-1; mutant; tumour; Cripto-3; CR-3;
XX cell proliferation.
XX
OS Homo sapiens.
XX
PN WO200222808-A2.
XX
PD 21-MAR-2002.
XX
PF 18-SEP-2001; 2001WO-US029066.
XX
PR 18-SEP-2000; 2000US-0233148P.
```

XX PA (BIOJ ) BIOGEN INC.  
XX PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
XX DR WPI; 2002-339868/37.  
XX PT New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful  
XX PT for treating cell proliferation, especially cancer, comprises amino acid  
XX PT change that prevents fucosylation at Thr88.  
XX PS Claim 2; Page 38; 41pp; English.  
XX CC The sequence represents a C-terminal truncated form of human Cripto-3.  
XX CC The invention relates to a novel mutant CRIPTO polypeptide, or its  
XX CC functional fragment, having at least one amino acid alteration at  
XX CC positions 86, 87 or 88. The mutant polypeptide, or its chimera, is used  
XX CC to inhibit growth of tumour cells, in vivo or in vitro, particularly for  
XX CC treating breast, ovarian, renal, colorectal, uterine, prostatic, lung,  
XX CC bladder or central nervous system cancers, melanoma and leukaemia, also  
XX CC generally for treating undesired cell proliferation  
XX SQ Sequence 129 AA;  
Query Match 100.0%; Score 96; DB 5; Length 129;  
Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 FRDSDSIWQEEPAPR 17  
Db 6 FRDSDSIWQEEPAPR 22  
RESULT 6  
ABB77103  
ID ABB77103 standard; protein; 139 AA.  
AC ABB77103;  
XX DT 08-OCT-2002 (first entry)  
XX DE Human Cripto-1 C-terminal truncated form (CR(cc)).  
XX KW Human; Cripto-1; CR-1; mutant; tumour; cell proliferation.  
XX OS Homo sapiens.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
XX FT Misc-difference 58 /note= "Wild type Thr substituted by Ala"  
XX FT  
XX PN WO200222808-A2.  
XX PD 21-MAR-2002.  
XX PF 18-SEP-2001; 2001WO-US029066.  
XX PR 18-SEP-2000; 2000US-0233148P.  
XX PA (BIOJ ) BIOGEN INC.  
XX PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
XX OS Homo sapiens.  
XX OS WO200222808-A2.  
XX PD 21-MAR-2002.  
XX PF 18-SEP-2001; 2001WO-US029066.  
XX PR 18-SEP-2000; 2000US-0233148P.  
XX PA (BIOJ ) BIOGEN INC.  
XX PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
XX DR WPI; 2002-339868/37.  
XX PT New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful  
XX PT for treating cell proliferation, especially cancer, comprises amino acid  
XX PT change that prevents fucosylation at Thr88.  
XX PS Claim 2; Page 35; 41pp; English.  
XX CC The sequence represents a C-terminal truncated form of human Cripto-1.  
XX CC The invention relates to a novel mutant CRIPTO polypeptide, or its  
XX CC functional fragment, having at least one amino acid alteration at  
XX CC positions 86, 87 or 88. The mutant polypeptide, or its chimera, is used  
XX CC to inhibit growth of tumour cells, in vivo or in vitro, particularly for  
XX CC treating breast, ovarian, renal, colorectal, uterine, prostatic, lung,

CC bladder or central nervous system cancers, melanoma and leukaemia, also  
CC generally for treating undesired cell proliferation  
XX SQ Sequence 139 AA;  
Query Match 100.0%; Score 96; DB 5; Length 139;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 FRDSDSIWQEEPAPR 17  
Db 16 FRDSDSIWQEEPAPR 32  
RESULT 7  
ABB77123  
ID ABB77123 standard; protein; 139 AA.  
XX AC ABB77123;  
XX DT 08-OCT-2002 (first entry)  
XX DE Human Cripto-3 C-terminal truncated form mutant CR-338(cc)T88A.  
XX KW Human; Cripto-1; CR-1; mutant; tumour; Cripto-3; CR-3;  
XX KW cell proliferation.  
XX OS Homo sapiens.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
XX FT Misc-difference 58 /note= "Wild type Thr substituted by Ala"  
XX FT  
XX PN WO200222808-A2.  
XX PD 21-MAR-2002.  
XX PF 18-SEP-2001; 2001WO-US029066.  
XX PR 18-SEP-2000; 2000US-0233148P.  
XX PA (BIOJ ) BIOGEN INC.  
XX PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
XX DR WPI; 2002-339868/37.  
XX PT New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful  
XX PT for treating cell proliferation, especially cancer, comprises amino acid  
XX PT change that prevents fucosylation at Thr88.  
XX PS Disclosure; Page 39; 41pp; English.  
XX CC The sequence represents a mutant C-terminal truncated form of human  
XX CC Cripto-3, where the threonine residue at position 58 has been replaced  
XX CC with an alanine residue. The invention relates to a novel mutant CRIPTO  
XX CC polypeptide, or its functional fragment, having at least one amino acid  
XX CC alteration at positions 86, 87 or 88. The mutant polypeptide, or its  
XX CC chimera, is used to inhibit growth of tumour cells, in vivo or in vitro,  
XX CC particularly for treating breast, ovarian, renal, colorectal, uterine,  
XX CC prostatic, lung, bladder or central nervous system cancers, melanoma and  
XX CC leukaemia, also generally for treating undesired cell proliferation  
XX SQ Sequence 139 AA;  
Query Match 100.0%; Score 96; DB 5; Length 139;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 FRDSDSIWQEEPAPR 17  
Db 16 FRDSDSIWQEEPAPR 32

```

RESULT 8
ABB77110
ID ABB77110 standard; protein; 139 AA.
AC ABB77110;
XX
XX
DT 08-OCT-2002 (first entry)
DE Human Cripto-1 C-terminal truncated form mutant CR38(cc)T88A.
XX
XX
KW Human; Cripto-1; CR-1; mutant; tumour; cell proliferation.
XX
XX
OS Homo sapiens.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 58
FT /note= "Wild type Thr substituted by Ala"
XX
XX
PN WO200222808-A2.
XX
XX
PD 21-MAR-2002.
XX
XX
PF 18-SEP-2001; 2001WO-US029066.
XX
XX
PR 18-SEP-2000; 2000US-0233148P.
XX
XX
PA (BIOJ ) BIOGEN INC.
XX
XX
PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;
XX
XX
DR WPI; 2002-339868/37.
XX
XX
PT New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful
PT for treating cell proliferation, especially cancer, comprises amino acid
PT change that prevents fucosylation at Thr88.
XX
XX
PS Example 3; Page 36; 41pp; English.
XX
XX
CC The sequence represents a mutant C-terminal truncated form of human
CC Cripto-1, where the threonine residue at position 58 has been replaced
CC with an alanine residue. The invention relates to a novel mutant CRIPTO
CC polypeptide, or its functional fragment, having at least one amino acid
CC alteration at positions 86, 87 or 88. The mutant polypeptide, or its
CC chimeric, is used to inhibit growth of tumour cells, in vivo or in vitro,
CC particularly for treating breast, ovarian, renal, colorectal, uterine,
CC prostatic, lung, bladder or central nervous system cancers, melanoma and
CC leukaemia, also generally for treating undesired cell proliferation
XX
XX
SQ Sequence 139 AA;

Query Match 100.0%; Score 96; DB 5; Length 139;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDINWPQEEPAIRPR 17
Db 16 FRDSDINWPQEEPAIRPR 32

RESULT 9
ABB77116
ID ABB77116 standard; protein; 139 AA.
AC ABB77116;
XX
XX
DT 08-OCT-2002 (first entry)
DE Human Cripto-3 C-terminal truncated form (CR-3(cc)).
XX
XX
KW Human; Cripto-1; CR-1; mutant; tumour; Cripto-3; CR-3;

```

```

KW cell proliferation.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200222808-A2.
XX
XX
PD 21-MAR-2002.
XX
XX
PF 18-SEP-2001; 2001WO-US029066.
XX
XX
PR 18-SEP-2000; 2000US-0233148P.
XX
XX
PA (BIOJ ) BIOGEN INC.
XX
XX
PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;
XX
XX
DR WPI; 2002-339868/37.
XX
XX
PT New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful
PT for treating cell proliferation, especially cancer, comprises amino acid
PT change that prevents fucosylation at Thr88.
XX
XX
PS Claim 2; Page 37; 41pp; English.
XX
XX
CC The sequence represents a C-terminal truncated form of human Cripto-3.
CC The invention relates to a novel mutant CRIPTO polypeptide, or its
CC functional fragment, having at least one amino acid alteration at
CC positions 86, 87 or 88. The mutant polypeptide, or its chimera, is used
CC to inhibit growth of tumour cells, in vivo or in vitro, particularly for
CC treating breast, ovarian, renal, colorectal, uterine, prostatic, lung,
CC bladder or central nervous system cancers, melanoma and leukaemia, also
CC generally for treating undesired cell proliferation
XX
XX
SQ Sequence 139 AA;

Query Match 100.0%; Score 96; DB 5; Length 139;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDINWPQEEPAIRPR 17
Db 16 FRDSDINWPQEEPAIRPR 32

RESULT 10
ADS88698
ID ADS88698 standard; protein; 173 AA.
XX
XX
AC ADS88698;
XX
XX
DT 16-DEC-2004 (first entry)
XX
XX
DE Amino acid sequence of a human secreted Cripto protein.
XX
XX
KW stem cell differentiation; cardiomyocyte; stem cell;
KW epidermal growth factor; EGF; EGF-CFC family; Cripto; neuronal cell;
KW Alq4; Nodal; heart disease; neuropathology; myocardial infarction;
KW Parkinson's disease; Alzheimer's disease; retinal degeneration;
KW cell therapy.
XX
XX
OS Homo sapiens.
XX
XX
PN WO2004083375-A2.
XX
XX
PD 30-SEP-2004.
XX
XX
PF 19-MAR-2004; 2004WO-IT000133.
XX
XX
PR 21-MAR-2003; 2003IT-RM000125.
XX
XX
PR 29-JUL-2003; 2003IT-RM000370.
XX
XX
PA (MINC/) MINCHIOTTI G.
XX
XX
PA (PERS/) PERSICO M.

```

PA (PARI//) PARISI S.  
XX  
XX Minchiotti G, Persico M, Parisi S;  
XX  
XX WPI; 2004-691035/67.  
XX  
XX Inducing stem cell differentiation in cardiomyocytes or neuronal cells by  
PT exposure to a protein of EGF-CFC or Cripto protein inhibitors,  
PT respectively, useful in treating heart diseases and neuropathologies.  
XX  
XX Disclosure; Page 13; 35pp; English.  
XX  
XX The specification describes a method for inducing stem cell  
CC differentiation into cardiomyocytes. The method comprises exposing the  
CC cells to a protein of the epidermal growth factor (EGF)-CFC family or its  
CC derivatives and having at least the EGF and CFC domains (e.g. Cripto  
CC protein). A second method is described for stem cell differentiation into  
CC neuronal cells, where the cells are exposed to an inhibitor of the Cripto  
CC protein. The Cripto protein inhibitor is an anti-Cripto antibody or its  
CC functional fragments, or is a peptide specifically selected from a random  
CC combinatorial peptide library, or is an antagonist of the Alq4(receptor)-  
CC Cripto(co-receptor)-Nodal(ligand) pathway. The differentiated stem cells  
CC of the invention are useful in the treatment of heart diseases and  
CC neuropathologies, e.g. myocardial infarction, Parkinson's disease,  
CC Alzheimer's disease and retinal degeneration. They are especially useful  
CC for cell therapy. The Cripto protein or its inhibitors is useful in the  
CC preparation of a composition able to direct stem cell differentiation  
CC toward the neuronal lineage. The present sequence represents a Cripto  
CC protein, which maybe used in the method of the invention.  
XX  
XX Sequence 173 AA;  
SQ  
  
Query Match 100.0%; Score 96; DB 8; Length 173;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 FRDDSIWQPQEPAIRPR 17  
Db 46 FRDDSIWQPQEPAIRPR 62  
|||||  
  
RESULT 11  
AAR13326  
ID AAR13326 standard; protein; 174 AA.  
XX  
XX AC AAR13326;  
XX  
XX 25-MAR-2003 (revised)  
DT 17-DEC-2001 (revised)  
DT 22-OCT-1991 (first entry)  
XX  
XX Recombinant CRIPTO protein.  
DE  
XX Cancer.  
XX  
XX Homo sapiens.  
OS  
XX USN7530165-N.  
FN  
XX 09-JUL-1991.  
PD  
XX 29-MAY-1990; 90US-00143529.  
PF  
XX 29-MAY-1990; 90US-00530165.  
PR  
XX (USSH ) NAT INST OF HEALTH.  
PA  
XX WPI; 1991-245692/33.  
DR  
XX New human CRIPTO gene - used for developing prods. for the study,  
PT diagnosis prognosis and treatment of human cancers.  
PT  
XX Disclosure; Fig 2; 24pp; English.  
PS

XX The sequence given is that of the recombinant E.coli derived CRIPTO  
CC protein, produced from CRIPTO cDNA. The availability of this protein make  
CC it possible to detect cells expressing the CRIPTO gene. It can be used to  
CC develop oligonucleotide probes, antisense oligonucleotides and antibodies  
CC for the study, diagnosis, prognosis and treatment of human cancer. The  
CC cDNA sequence has been deposited in the EMBL Gene Data Bank. This  
CC sequence and the amino acid sequence of the corresponding natural CRIPTO  
CC protein (different to the recombinant sequence) are given in the  
CC specification but poor print quality prevents their inclusion in this  
CC database. (Note: Revised entry submitted to correct the patent number  
CC format of US Government-owned NTIS applications to prevent clashes with  
CC ongoing US granted patent numbers. For further information please visit  
CC the Derwent web site at [www.derwent.com/dwpi/updates/ntis\\_us.html](http://www.derwent.com/dwpi/updates/ntis_us.html).)  
CC (Updated on 25-MAR-2003 to correct PR field.)  
XX  
XX Sequence 174 AA;  
SQ  
  
Query Match 100.0%; Score 96; DB 2; Length 174;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 FRDDSIWQPQEPAIRPR 17  
Db 32 FRDDSIWQPQEPAIRPR 48  
|||||  
  
RESULT 12  
AAW32107  
ID AAW32107 standard; protein; 174 AA.  
XX  
XX AC AAW32107;  
XX  
XX 25-MAR-2003 (revised)  
DT 17-FEB-1998 (first entry)  
DT  
XX Recombinant human CRIPTO protein expressed in E. coli.  
DE  
XX CRIPTO gene; human; colorectal carcinoma; breast carcinoma; cancer;  
KW tumour; diagnosis; prognosis; therapy; antibody; Escherichia coli.  
XX  
XX Homo sapiens.  
OS  
XX Synthetic.  
OS  
XX US5654140-A.  
FN  
XX 05-AUG-1997.  
PD  
XX 10-NOV-1994; 94US-00337911.  
PF  
XX 29-MAY-1990; 90US-00530165.  
PR  
XX 18-SEP-1992; 92US-00947315.  
PR  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA  
XX Salomon DS, Persico MG;  
PI  
XX WPI; 1997-401840/37.  
DR  
XX New cloned CRIPTO gene - used for diagnosis and prognosis of carcinoma,  
PT also anti-sense sequences for tumour treatment.  
PT  
XX Disclosure; Fig 2; 13pp; English.  
PS  
XX This protein comprises recombinant human CRIPTO produced in Escherichia  
CC coli. Human CRIPTO cDNA (see AAT88973) was expressed in E. coli inclusion  
CC bodies as a 20-22 kDa protein. CRIPTO is a member of the epidermal growth  
CC factor family. Elevated levels of CRIPTO mRNA, or of protein translated  
CC from it, are diagnostic and prognostic of carcinoma, especially  
CC colorectal and breast carcinoma. Recombinant CRIPTO can be expressed in  
CC prokaryotic or eukaryotic host cells, and used for the identification and  
CC characterisation of specific cell surface receptors, or to raise specific  
CC antibodies to assay CRIPTO protein expression in immunoassays. (Updated

```
CC on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 174 AA;

Query Match      100.0%; Score 96; DB 2; Length 174;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17
   |||||
Db 32 FRDSDIWPQEEPAIRPR 48

RESULT 13
ADO05065
ID ADO05065 standard; protein; 174 AA.
XX
AC ADO05065;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human cripto growth factor protein.
XX
KW Criptin growth factor; CGF; wound healing; tissue regeneration;
KW implant fixation; angiogenesis; neoplasia; tumour; gene therapy; human;
KW cripto growth factor.
XX
OS Homo sapiens.
XX
PN US2004086967-A1.
XX
PD 06-MAY-2004.
XX
PF 22-SEP-2003; 2003US-00665602.
XX
PR 06-JUN-1995; 95US-00471371.
XX
PR 09-SEP-1999; 99US-00393023.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Meissner PS, Coleman TA;
XX
DR WPI; 2004-356201/33.
XX
PT New human polynucleotides encoding human criptin growth factor
PT polypeptides, useful for wound healing or tissue regeneration.
PT stimulating implant fixation and angiogenesis, and for treating and/or
PT preventing tumor.
XX
PS Disclosure; SEQ ID NO 7; 19pp; English.
XX
CC The invention provides criptin growth factor (CGF) polypeptides and their
CC encoding polynucleotides. The invention is useful for wound healing and
CC tissue regeneration, stimulating implant fixation, angiogenesis and for
CC treating and preventing neoplasia such as tumour. The invention is also
CC useful in gene therapy. The present sequence is human cripto growth
CC factor protein. This sequence is used in the invention.
XX
SQ Sequence 174 AA;

Query Match      100.0%; Score 96; DB 8; Length 174;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17
   |||||
Db 32 FRDSDIWPQEEPAIRPR 48

RESULT 14
ADS88699
ID ADS88699 standard; protein; 183 AA.
XX
```

---

```
AC
XX
DT 16-DEC-2004 (first entry)
XX
DE Amino acid sequence of a human secreted His-tagged Cripto protein.
XX
KW stem cell differentiation; cardiomyocyte; stem cell;
KW epidermal growth factor; EGF; EGF-CFC family; Cripto; neuronal cell;
KW Alq4; Nodal; heart disease; neuropathology; myocardial infarction;
KW Parkinson's disease; Alzheimer's disease; retinal degeneration;
KW cell therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 174..183
FT /note= "His tag"
XX
PN WO2004083375-A2.
XX
PD 30-SEP-2004.
XX
PF 19-MAR-2004; 2004WO-IT000133.
XX
PR 21-MAR-2003; 2003IT-RM000125.
PR 29-JUL-2003; 2003IT-RM000370.
XX
XX (MING/) MINCHIOTTI G.
PA (PERS/) PERSICO M.
PA (PARI/) PARISI S.
XX
PI Minchiotti G, Persico M, Parisi S;
XX
PF WPI; 2004-691035/67.
XX
PT Inducing stem cell differentiation in cardiomyocytes or neuronal cells by
PT exposure to a protein of EGF-CFC or Cripto protein inhibitors,
PT respectively, useful in treating heart diseases and neuropathologies.
XX
PS Disclosure; Page 13; 35pp; English.
XX
CC The specification describes a method for inducing stem cell
CC differentiation into cardiomyocytes. The method comprises exposing the
CC cells to a protein of the epidermal growth factor (EGF)-CFC family or its
CC derivatives and having at least the EGF and CFC domains (e.g. Cripto
CC protein). A second method is described for stem cell differentiation into
CC neuronal cells, where the cells are exposed to an inhibitor of the Cripto
CC protein. The Cripto protein inhibitor is an anti-Cripto antibody or its
CC functional fragments, or is a peptide specifically selected from a random
CC combinatorial peptide library, or is an antagonist of the Alq4(receptor)-
CC Cripto(co-receptor)-Nodal(ligand) pathway. The differentiated stem cells
CC of the invention are useful in the treatment of heart diseases and
CC neuropathologies, e.g. myocardial infarction, Parkinson's disease,
CC Alzheimer's disease and retinal degeneration. They are especially useful
CC for cell therapy. The Cripto protein or its inhibitors is useful in the
CC preparation of a composition able to direct stem cell differentiation
CC toward the neuronal lineage. The present sequence represents a Cripto
CC protein, which maybe used in the method of the invention.
XX
SQ Sequence 183 AA;

Query Match      100.0%; Score 96; DB 8; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17
   |||||
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 15
ADO28629
```

AD028629 standard; protein; 184 AA.  
ADO28629;  
12-AUG-2004 (first entry)  
Human CRIPTO CR-1 protein SEQ ID NO:58.  
high-grade dysplasia; HGD; oesophageal adenocarcinoma;  
neo-plastic transformation; cancer; cytostatic; gene therapy; human;  
CRIPTO CR-1; chromosome 3.  
Homo sapiens.  
WO2004044178-A2.  
27-MAY-2004.  
13-NOV-2003; 2003WO-US036260.  
13-NOV-2002; 2002US-0425813P.  
(GETH ) GENENTECH INC.  
Smith V;  
WPI; 2004-420319/39.  
N-PSDB; ADO28628.  
Detecting of high-grade dysplasia in cells of a mammalian tissue sample  
comprises establishing the level of expression in the test tissue sample  
of the genes.  
Example 4; SEQ ID NO 58; 256pp; English.  
The present invention describes a method for detecting high-grade  
dysplasia (HGD) in cells of a mammalian tissue sample. Also described:  
(1) identifying an oesophageal tissue susceptible to oesophageal  
adenocarcinoma; (2) determining the predisposition of a mammalian tissue  
to a neo-plastic transformation by detecting HGD in cells of the tissue;  
and (3) detecting cancer in a patient. The method can be used in  
detecting HGD and cancer in cells of a mammalian tissue sample. The  
methods and compositions of the present invention can be used in treating  
and preventing HGD and cancer, and in gene therapy. The present sequence  
represents human CRIPTO CR-1, which is used in the exemplification of the  
present invention. The human CRIPTO CR-1 gene is located on chromosome 3.  
Sequence 184 AA;  
Query Match 100.0%; Score 96; DB 8; Length 184;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FRDSDIWQEEPAIRPR 17  
DB 44 FRDSDIWQEEPAIRPR 60  
RESULT 16  
AAE36467  
ID AAE36467 standard; protein; 187 AA.  
XX  
AC AAE36467;  
XX  
DT 07-AUG-2003 (first entry)  
XX  
DE Human colon tumour protein from clone, C1584P.  
XX  
KW Human; colon tumour protein; cancer; gene therapy; vaccine; TDGF1;  
KW teratocarcinoma-derived growth factor 1.  
XX  
OS Homo sapiens.  
XX  
PN WO2003022126-A2.  
XX  
PD 20-MAR-2003.  
XX  
PF 20-AUG-2002; 2002WO-US026601.  
XX  
PR 21-AUG-2001; 2001US-0314221P.  
XX  
PR 21-DEC-2001; 2001US-0343517P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
XX Meagher MJ, King GE, Secrist H, Jiang Y, Stolk JA;  
PI WPI; 2003-300990/29.  
XX  
DR N-PSDB; AAD55563.  
XX  
XX New isolated polynucleotide useful for stimulating and/or expanding T-  
PT cells specific for a tumor protein, and preventing, diagnosing and/or  
PT treating cancer, particularly colon cancer, in patients.  
XX  
PS Claim 2; Page 156; 158pp; English.  
XX  
CC The invention relates to compositions and methods for the diagnosis and  
CC therapy of cancer, particularly colon cancer. The invention also provides  
CC colon tumour proteins and to polynucleotides encoding such proteins.  
CC Polynucleotides of the invention are useful for preventing, diagnosing  
CC and/or treating cancer, particularly colon cancer in a patient. They are  
CC used in gene therapy and also to prepare vaccines. The present sequence  
CC is human colon tumour protein from clone C1584P also known as  
CC teratocarcinoma-derived growth factor 1 (TDGF1)  
XX  
SQ Sequence 187 AA;  
Query Match 100.0%; Score 96; DB 6; Length 187;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FRDSDIWQEEPAIRPR 17  
DB 46 FRDSDIWQEEPAIRPR 62  
RESULT 17  
AAE22548  
ID AAR22548 standard; protein; 188 AA.  
XX  
AC AAR22548;  
XX  
XX 25-MAR-2003 (revised)  
DT 17-DEC-2001 (revised)  
DT 28-AUG-1992 (first entry)  
XX  
DE Human CRIPTO CR-1 protein.  
XX  
KW cell proliferation; tumour; CR-3; transforming growth factor;  
KW epidermal growth factor.  
XX  
OS Homo sapiens.  
XX  
PN USN7749001-N.  
XX  
PD 25-FEB-1992.  
XX  
PF 23-AUG-1991; 91US-00749001.  
XX  
PR 23-AUG-1991; 91US-00749001.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
XX  
PI Salomon D, Persico M;  
XX  
DR WPI; 1992-123675/15.  
DR N-PSDB; AAQ22495.



XX New crypto gene CR-1 and cripto-related gene CR-3 genomic DNA - CR-3  
PT protein and anti-CR-3 antibodies, useful in immunoassay to detect CR-3 as  
PT tumour specific marker.

XX Disclosure; Page 22; 44pp; English.

XX The CR-1 genomic clone which codes for this protein can be used in  
CC transgenic animals to examine the effects of overexpression of this gene  
CC on development and tumorigenicity and to study regulation of CR-1 gene.  
CC See also AAQ22494. (Note: Revised entry submitted to correct the patent  
CC number format of US Government-owned NTIS applications to prevent clashes  
CC with ongoing US granted patent numbers. For further information please  
CC visit the Derwent web site at [www.derwent.com/dwpi/updates/ntis\\_us.html](http://www.derwent.com/dwpi/updates/ntis_us.html).)  
CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 188 AA;

Query Match 100.0%; Score 96; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FRDSDIWPQEEPAIRPR 17  
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 18  
AAR22547  
ID AAR22547 standard; protein; 188 AA.

AC AAR22547;

XX 25-MAR-2003 (revised)  
DT 17-DEC-2001 (revised)  
DT 28-AUG-1992 (first entry)

DE Human CRIPTO-related protein CR-3.

XX cell proliferation; tumour; CR-1; transforming growth factor;  
KW epidermal growth factor.

XX Homo sapiens.

XX USN7749001-N.

XX 25-FEB-1992.

XX 23-AUG-1991; 91US-00749001.

XX 23-AUG-1991; 91US-00749001.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.

XX Salomon D, Persico M;

XX WPI; 1992-123675/15.

XX N-PSDB; AAQ22494.

XX New crypto gene CR-1 and cripto-related gene CR-3 genomic DNA - CR-3  
PT protein and anti-CR-3 antibodies, useful in immunoassay to detect CR-3 as  
PT tumour specific marker.

XX Disclosure; Page 33; 44pp; English.

XX The CR-3 nucleotide sequence was isolated from a human genomic library  
CC using a labelled DNA fragment containing 800bp upstream of the  
CC translation initiation site of CR-1. Clones were isolated which contained  
CC the 5' cDNA non-coding region of CR-3. One clone was sequenced and found  
CC to include a complete CRIPTO cDNA lacking introns and containing a  
CC poly(A) tract at the 3' end. Seven single base pair substitutions were  
CC observed in the coding region (c.f. CR-1 sequence in *Clecodicola A.* et  
CC al., EMBO J.8:1987-1991(1989)), six of which gave rise to amino acid

CC changes. The changes occur at the following amino acid positions: 6, 22,  
CC 43, 68, 86 and 178. (Note: Revised entry submitted to correct the patent  
CC number format of US Government-owned NTIS applications to prevent clashes  
CC with ongoing US granted patent numbers. For further information please  
CC visit the Derwent web site at [www.derwent.com/dwpi/updates/ntis\\_us.html](http://www.derwent.com/dwpi/updates/ntis_us.html).)  
CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 188 AA;

Query Match 100.0%; Score 96; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FRDSDIWPQEEPAIRPR 17  
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 19  
AAW25667  
ID AAW25667 standard; protein; 188 AA.

XX AC AAW25667;

XX 25-MAR-2003 (revised)  
DT 04-NOV-1997 (first entry)

DE Protein encoded by CRIPTO-related gene, CR-3.

XX CRIPTO-related gene; CR-3; epidermal growth factor; EGF; TGF-alpha;  
KW amphiregulin; tumour specific marker; colon cancer cell line;  
KW colorectal tumour; mesenchyme; epithelial cell.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 7 /note= "Ala>Val mutation"

FT Misc-difference 22 /note= "Val>Ala mutation"

FT Misc-difference 43 /note= "Tyr>Asp mutation"

FT Misc-difference 68 /note= "Pro>Leu mutation"

FT Misc-difference 86 /note= "Gly>Glu mutation"

FT Misc-difference 178 /note= "Val>Ala mutation"

XX US5650285-A.

XX 22-JUL-1997.

XX 05-JUN-1995; 95US-00463335.

XX 23-AUG-1991; 91US-00749001.

XX 17-NOV-1993; 93US-00154198.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Persica MG, Salomon DS;

XX WPI; 1997-384668/35.

XX N-PSDB; AAT80987.

XX Assays for CRIPTO-related gene product CR-3 - used in studies on the  
PT regulation of the proliferation, differentiation and transformation of  
PT cells.

XX Claim 2; Fig 2; 25pp; English.

XX This sequence is encoded by the CRIPTO-related gene, CR-3. The CR-3 gene  
CC sequence is identical to the human CRIPTO gene with the exception of

CC eight base pair substitutions that give rise to six amino acid changes.  
CC CR-3 exhibits partial amino acid sequence homology and a tertiary  
CC structure within a 38 amino acid region similar to the epidermal growth  
CC factor (EGF) supergens family that includes EGF, TGF-alpha and  
CC amphiregulin. Expression of CR-3 may serve as a tumour specific marker as  
CC it is expressed in several human colon cancer cell lines and possibly in  
CC human colorectal tumours. The assay of the amount of CR-3 in a sample can  
CC be used in studies on the regulation of the proliferation,  
CC differentiation, and transformation of various mesenchymal and epithelial  
CC cells. Study of the expression of CR-3 may also be used in the diagnosis,  
CC prognosis and treatment of tumours. (Updated on 25-MAR-2003 to correct PF  
CC field.)  
XX  
SQ

Query Match 100.0%; Score 96; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDDSIWQEPAPRPR 17  
| | | | | | | | | | | | | | | | | |  
Db 46 FRDDSIWQEPAPRPR 62

RESULT 20  
AAW32108  
ID AAW32108 standard; protein; 188 AA.  
AC AAW32108;  
XX  
XX  
DT 25-MAR-2003 (revised)  
DT 17-FEB-1998 (first entry)  
XX  
XX  
DE Human CRIPTO protein.  
XX  
KW CRIPTO gene; human; colorectal carcinoma; breast carcinoma; cancer;  
KW tumour; diagnosis; prognosis; therapy; antibody.  
XX  
XX Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT Cleavage-site 28. .29  
FT Misc-difference 38  
FT /note= "encoded by CCT"  
FT Modified-site 79. .81  
FT /note= "Asn is N-glycosylated"  
FT Misc-difference 87  
FT /note= "encoded by GGA"  
FT Cleavage-site 126. .127  
FT Misc-difference 151  
FT /note= "encoded by GGC"  
FT Cleavage-site 171. .174  
XX  
XX US5654140-A.  
PN  
XX  
XX 05-AUG-1997.  
PD  
XX  
XX 10-NOV-1994; 94US-00337911.  
PF  
XX  
XX 29-MAY-1990; 90US-00530165.  
PR  
XX 18-SEP-1992; 92US-00947315.  
PR  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA  
XX Salomon DS, Persico MG;  
XX  
XX WPI; 1997-401840/37.  
DR  
XX N-PSDB; AAT88973.  
DR  
XX New cloned CRIPTO gene - used for diagnosis and prognosis of carcinoma,  
PT also anti-sense sequences for tumour treatment.  
PT  
XX  
XX Disclosure; Fig 1A-B; 13pp; English.  
PS

XX This protein comprises human CRIPTO, a protein encoded by the novel  
CC CRIPTO gene that is associated with cancers such as colorectal cancer and  
CC breast carcinoma. Its sequence was deduced from a CRIPTO cDNA clone (see  
CC AAT88973) obtained from human teratocarcinoma cell line NT2D1. It is a  
CC member of the epidermal growth factor family. Elevated levels of CRIPTO  
CC mRNA, or of protein translated from it, are diagnostic and prognostic of  
CC carcinoma, especially colorectal and breast carcinoma. Recombinant CRIPTO  
CC can be expressed in prokaryotic or eukaryotic host cells (see AAW32107),  
CC and used for the identification and characterisation of specific cell  
CC surface receptors, or to raise specific antibodies to assay CRIPTO  
CC protein expression in immunoassays. (Updated on 25-MAR-2003 to correct PF  
CC field.)  
XX  
SQ

Query Match 100.0%; Score 96; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDDSIWQEPAPRPR 17  
| | | | | | | | | | | | | | | | | |  
Db 46 FRDDSIWQEPAPRPR 62

RESULT 21  
AAW19980  
ID AAW19980 standard; protein; 188 AA.  
AC AAW19980;  
XX  
XX  
DT 25-MAR-2003 (revised)  
DT 19-AUG-1997 (first entry)  
XX  
XX  
DE Human CRIPTO-related gene product, CR-3, a tumour-marker.  
XX  
KW CRIPTO; CR1; CR3; epidermal growth factor superfamily; EGF; mitogen;  
KW tumour marker; epithelium; mesenchyme; diagnosis; prognosis.  
XX  
XX Homo sapiens.  
XX  
XX US5620866-A.  
PN  
XX  
XX 15-APR-1997.  
PD  
XX  
XX 17-NOV-1993; 93US-00154198.  
PF  
XX  
XX 23-AUG-1991; 91US-00749001.  
PR  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA  
XX Salomon DS, Persico MG;  
XX  
XX WPI; 1997-235180/21.  
DR  
XX N-PSDB; AAT71145.  
DR  
XX New DNA encoding the human CRIPTO-related gene product CR-3 - used to  
PT produce recombinant protein, useful for preparation of diagnostic  
PT antibodies.  
XX  
XX Claim 2; Col 21-22; 25pp; English.  
PS  
XX AAW19980 shows the amino acid (aa) sequence of human CRIPTO-related gene  
CC product CR-3. CR-3 is thought to be a tumour-specific marker and as such  
CC DNA and aa sequences of CR-3 and derivatives can be used in the  
CC diagnosis, prognosis and possibly treatment of some types of cancer. CR-3  
CC includes a region homologous to products of the epidermal growth factor  
CC superfamily of genes and it may be a mitogen involved in regulating  
CC proliferation, differentiation and transformation of mesenchymal and  
CC epithelial cells. (Updated on 25-MAR-2003 to correct PF field.)  
XX  
SQ

Query Match 100.0%; Score 96; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17  
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 22  
AAW29735  
ID AAW29735 standard; protein; 188 AA.

XX AC AAW29735;

XX DT 25-MAR-2003 (revised)  
XX DT 09-NOV-1998 (first entry)

XX DE Homo sapiens CRIPTO protein.

XX KW CRIPTO; diagnosis; prognosis; cancer; colorectal; breast.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Cleavage-site 28..29

FT /note= "V-A"

FT Region 79..81

FT /note= "potential Asp glycosylation sequence"

FT Cleavage-site 111..112

FT /note= "R-K"

FT Cleavage-site 126..127

FT /note= "K-K"

FT Cleavage-site 159..160

FT /note= "V-A"

FT Cleavage-site 171..174

FT /note= "R-T-T-T"

XX US5792616-A.

XX PD 11-AUG-1998.

XX PF 05-JUN-1995; 95US-00463616.

XX PR 28-MAY-1990; 90US-00530165.

XX PR 28-SEP-1992; 92US-00947315.

XX PR 10-NOV-1994; 94US-00337911.

XX PA (USGO ) US GOVERNMENT.

XX PI Persico MG, Salomon DS;

XX DR WPI; 1998-456123/39.

XX DR N-PSDB; AAV47553.

XX PT New antibody which binds a CRIPTO protein - is useful for screening for  
XX PT expression of a CRIPTO protein in a tissue sample.

XX PS Disclosure; Fig 2; 12pp; English.

XX CC The sequence is that of the human CRIPTO protein. Expression of CRIPTO  
XX CC mRNA and CRIPTO protein appears to be associated with various cancers,  
XX CC e.g. colon, breast and small cell lung carcinomas and thus the sequence  
XX CC or antibodies raised against it may be of use in the diagnosis of  
XX CC cancers. (Updated on 25-MAR-2003 to correct PF field.)

XX SQ Sequence 188 AA;

Query Match 100.0%; Score 96; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17

Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 23  
AAW87631

ID AAW87631 standard; protein; 188 AA.

XX AC AAW87631;

XX DT 03-MAR-1999 (first entry)

XX DE Human CRIPTO-related (CR-3) protein.

XX KW human CRIPTO-related protein; CR-1; antibody; CR-3; proliferation;  
XX KW differentiation; transformation; mesenchymal cell; epithelial cell;  
XX KW tumour specific marker; cancer.

XX OS Homo sapiens.

XX PN US5854399-A.

XX PD 29-DEC-1998.

XX PF 05-JUN-1995; 95US-00464023.

XX PR 23-AUG-1991; 91US-00749001.

XX PR 17-NOV-1993; 93US-00154198.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Persica MG, Salomon DS;

XX DR WPI; 1999-095001/08.

XX DR N-PSDB; AAV83924.

XX PT New antibody to human CRIPTO-related polypeptide-3 - used for the  
XX PT detection of CRIPTO-related polypeptide-3 which is involved in the  
XX PT regulation of the proliferation, differentiation and transformation of  
XX PT cells.

XX PS Claim 1; Col 21-22; 26pp; English.

XX CC The present sequence represents a human CRIPTO-related protein CR-3. The  
XX CC specification describes an antibody which has binding affinity to CR-3  
XX CC and not to CR-1. The antibodies can be used for the detection of CR-3. CR  
XX CC -3 is a regulatory molecule involved in regulating the proliferation,  
XX CC differentiation, and transformation of various mesenchymal and epithelial  
XX CC cells. In addition expression of CR-3 may serve as a tumour specific  
XX CC marker that may have applicability in the diagnosis, prognosis and  
XX CC possible treatment of specific types of cancer

XX SQ Sequence 188 AA;

Query Match 100.0%; Score 96; DB 2; Length 188;

Best Local Similarity 100.0%; Pred. No. 1.6e-06;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17

Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 24  
AAW87630

ID AAW87630 standard; protein; 188 AA.

XX AC AAW87630;

XX DT 03-MAR-1999 (first entry)

XX DE Human CRIPTO-related (CR-1) protein.

KW human CRIPTO-related protein; CR-1; antibody; CR-3; proliferation;  
KW differentiation; transfection; mesenchymal cell; epithelial cell;  
KW tumour specific marker; cancer.  
XX Homo sapiens.  
XX  
XX US5854399-A.  
XX  
XX 29-DEC-1998.  
XX  
XX 05-JUN-1995; 95US-00464023.  
XX  
XX 23-AUG-1991; 91US-00749001.  
XX 17-NOV-1993; 93US-00154198.  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Persica MG, Salomon DS;  
XX WPI; 1999-095001/08.  
XX N-PSDB; AAV83923.  
XX  
XX New antibody to human CRIPTO-related polypeptide-3 - used for the  
PT detection of CRIPTO-related polypeptide-3 which is involved in the  
PT regulation of the proliferation, differentiation and transformation of  
PT cells.  
XX  
XX Example 2; Col 17-18; 26pp; English.  
XX  
XX The present sequence represents a human CRIPTO-related protein CR-1. The  
CC specification describes an antibody which has binding affinity to CR-3  
CC and not to CR-1. The antibodies can be used for the detection of CR-3. CR  
CC -3 is a regulatory molecule involved in regulating the proliferation,  
CC differentiation, and transformation of various mesenchymal and epithelial  
CC cells. In addition expression of CR-3 may serve as a tumour specific  
CC marker that may have applicability in the diagnosis, prognosis and  
CC possible treatment of specific types of cancer  
XX  
XX Sequence 188 AA;  
SQ  
Query Match 100.0%; Score 96; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FRDSDIWPQEPFAIRPR 17  
DB 46 FRDSDIWPQEPFAIRPR 62  
RESULT 25  
ABB77102  
ID ABB77102 standard; protein; 188 AA.  
XX  
XX ABB77102;  
XX  
XX 08-OCT-2002 (first entry)  
XX  
XX Human CRIPTO-3 full length protein.  
XX  
XX Human; CRIPTO-1; CR-1; mutant; tumour; CRIPTO-3; CR-3;  
KW cell proliferation.  
XX  
XX Homo sapiens.  
XX  
XX WO200222808-A2.  
XX  
XX 21-MAR-2002.  
XX  
XX 18-SEP-2001; 2001WO-US029066.  
XX  
XX 18-SEP-2000; 2000US-0233148P.  
XX  
XX (BIOI ) BIOGEN INC.  
PA

XX Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
PI WPI; 2002-339868/37.  
XX N-PSDB; ABB55853.  
DR  
XX  
XX New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful  
PT for treating cell proliferation, especially cancer, comprises amino acid  
PT change that prevents fucosylation at Thr88.  
XX  
XX Claim 2; Fig 1; 41pp; English.  
XX  
XX The sequence represents the full length human CRIPTO-3 protein. The  
CC invention relates to a novel mutant CRIPTO polypeptide, or its functional  
CC fragment, having at least one amino acid alteration at positions 86, 87  
CC or 88. The mutant polypeptide, or its chimera, is used to inhibit growth  
CC of tumour cells, in vivo or in vitro, particularly for treating breast,  
CC ovarian, renal, colorectal, uterine, prostatic, lung, bladder or central  
CC nervous system cancers, melanoma and leukaemia, also generally for  
CC treating undesired cell proliferation  
XX  
XX Sequence 188 AA;  
SQ  
Query Match 100.0%; Score 96; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FRDSDIWPQEPFAIRPR 17  
DB 46 FRDSDIWPQEPFAIRPR 62  
RESULT 26  
ABB77101  
ID ABB77101 standard; protein; 188 AA.  
XX  
XX ABB77101;  
XX  
XX 08-OCT-2002 (first entry)  
XX  
XX Human CRIPTO-1 full length protein.  
XX  
XX Human; CRIPTO-1; CR-1; mutant; tumour; cell proliferation.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Peptide 1..30  
FT /label= Signal\_peptide  
FT Protein 31..188  
FT /label= Mature\_Cripto-1  
XX  
XX WO200222808-A2.  
XX  
XX 21-MAR-2002.  
XX  
XX 18-SEP-2001; 2001WO-US029066.  
XX  
XX 18-SEP-2000; 2000US-0233148P.  
XX  
XX (BIOI ) BIOGEN INC.  
XX  
XX Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
PI WPI; 2002-339868/37.  
XX N-PSDB; ABL55852.  
XX  
XX New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful  
PT for treating cell proliferation, especially cancer, comprises amino acid  
PT change that prevents fucosylation at Thr88.  
XX  
XX Claim 2; Fig 1; 41pp; English.  
XX  
XX

CC The sequence represents the full length human Cripto-1 protein. The  
CC invention relates to a novel mutant CRIPTO polypeptide, or its functional  
CC fragment, having at least one amino acid alteration at positions 86, 87  
CC or 88. The mutant polypeptide, or its chimera, is used to inhibit growth  
CC of tumour cells, in vivo or in vitro, particularly for treating breast,  
CC ovarian, renal, colorectal, uterine, prostatic, lung, bladder or central  
CC nervous system cancers, melanoma and leukaemia, also generally for  
CC treating undesired cell proliferation  
XX Sequence 188 AA;

Query Match 100.0%; Score 96; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17  
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 27  
ABB77122  
ID ABB77122 standard; protein; 188 AA.

AC ABB77122;

XX 08-OCT-2002 (first entry)

XX Human Cripto-3 full length protein mutant CR-3T88A.

DE Human; Cripto-1; CR-1; mutant; tumour; Cripto-3; CR-3;  
XX cell proliferation.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 88 /note= "Wild type Thr substituted by Ala"

XX WO200222808-A2.

XX 21-MAR-2002.

XX 18-SEP-2001; 2001WO-US029066.

XX 18-SEP-2000; 2000US-0233148P.

XX (BIOJ ) BIOGEN INC.

XX Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
XX WPI; 2002-339868/37.

XX New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful  
XX for treating cell proliferation, especially cancer, comprises amino acid  
XX change that prevents fucosylation at Thr88.

XX Disclosure; Page 39; 41pp; English.

XX The sequence represents the full length human Cripto-3 protein where the  
XX threonine residue at position 88 has been replaced with an alanine  
XX residue. The invention relates to a novel mutant CRIPTO polypeptide, or  
XX its functional fragment, having at least one amino acid alteration at  
XX positions 86, 87 or 88. The mutant polypeptide, or its chimera, is used  
XX to inhibit growth of tumour cells, in vivo or in vitro, particularly for  
XX treating breast, ovarian, renal, colorectal, uterine, prostatic, lung,  
XX bladder or central nervous system cancers, melanoma and leukaemia, also  
XX generally for treating undesired cell proliferation

XX Sequence 188 AA;

Query Match 100.0%; Score 96; DB 5; Length 188;

Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17  
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 28

ABB77109  
ID ABB77109 standard; protein; 188 AA.

XX ABB77109;

XX 08-OCT-2002 (first entry)

XX Human Cripto-1 full length protein mutant CRT88A.

XX Human; Cripto-1; CR-1; mutant; tumour; cell proliferation.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 88 /note= "Wild type Thr substituted by Ala"

XX WO200222808-A2.

XX 21-MAR-2002.

XX 18-SEP-2001; 2001WO-US029066.

XX 18-SEP-2000; 2000US-0233148P.

XX (BIOJ ) BIOGEN INC.

XX Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
XX WPI; 2002-339868/37.

XX New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful  
XX for treating cell proliferation, especially cancer, comprises amino acid  
XX change that prevents fucosylation at Thr88.

XX Example 3; Page 36; 41pp; English.

XX The sequence represents the full length human Cripto-1 protein where the  
XX threonine residue at position 88 has been replaced with an alanine  
XX residue. The invention relates to a novel mutant CRIPTO polypeptide, or  
XX its functional fragment, having at least one amino acid alteration at  
XX positions 86, 87 or 88. The mutant polypeptide, or its chimera, is used  
XX to inhibit growth of tumour cells, in vivo or in vitro, particularly for  
XX treating breast, ovarian, renal, colorectal, uterine, prostatic, lung,  
XX bladder or central nervous system cancers, melanoma and leukaemia, also  
XX generally for treating undesired cell proliferation

XX Sequence 188 AA;

Query Match 100.0%; Score 96; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17  
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 29

AAO14637  
ID AAO14637 standard; protein; 188 AA.

XX AAO14637;

```
XX 13-JUN-2002 (first entry)
XX Human cripto 3 protein.
DE Human; cripto 3 protein; epidermal growth factor family; EGF family;
XX gene therapy; protein therapy; vaccine; lung cancer; colon cancer;
XX colorectal cancer; breast cancer.
XX Homo sapiens.
OS WO200216413-A2.
XX FN
XX PD
XX 28-FEB-2002.
XX 20-AUG-2001; 2001WO-EP009646.
XX 24-AUG-2000; 2000GB-00020953.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX Cassart J, Coche T, Palmantier RM, Vinals Y De BassolsC;
XX WPI; 2002-280910/32.
XX N-PSDB; AAL42197.
XX New Cripto tumor polynucleotides, useful in medicine, particularly for
XX treating tumors (e.g. lung, colorectal, colon or breast tumor) that
XX express a cripto antigen.
XX Claim 5; Page 94; 134pp; English.
XX The invention comprises the nucleotide and amino acid sequences of human
XX cripto proteins. Cripto is a 188 amino acid protein that shares
XX homologues with the epidermal growth factor (EGF) family. The cripto
XX nucleic acid and protein sequences of the invention are useful for
XX treating (e.g. gene therapy, protein therapy, vaccine) cancer, such as:
XX lung, colon, colorectal or breast. The present amino acid sequence
XX represents the human cripto 3 protein
XX Sequence 188 AA;
XX Query Match 100.0%; Score 96; DB 5; Length 188;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-06;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 FRDDSIWQPQEEPAIRPR 17
XX 46 FRDDSIWQPQEEPAIRPR 62
XX
XX RESULT 30
XX AAO14638
XX ID AAO14638 standard; protein; 188 AA.
XX AC AAO14638;
XX 13-JUN-2002 (first entry)
XX Human cripto protein.
XX Human; cripto protein; epidermal growth factor family; EGF family;
XX gene therapy; protein therapy; vaccine; lung cancer; colon cancer;
XX colorectal cancer; breast cancer.
XX Homo sapiens.
OS WO200216413-A2.
XX FN
XX PD
XX 28-FEB-2002.
XX 20-AUG-2001; 2001WO-EP009646.
XX
```

```
PR 24-AUG-2000; 2000GB-00020953.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX Cassart J, Coche T, Palmantier RM, Vinals Y De BassolsC;
XX WPI; 2002-280910/32.
XX N-PSDB; AAL42198.
XX New Cripto tumor polynucleotides, useful in medicine, particularly for
XX treating tumors (e.g. lung, colorectal, colon or breast tumor) that
XX express a cripto antigen.
XX Disclosure; Page 95; 134pp; English.
XX The invention comprises the nucleotide and amino acid sequences of human
XX cripto proteins. Cripto is a 188 amino acid protein that shares
XX homologues with the epidermal growth factor (EGF) family. The cripto
XX nucleic acid and protein sequences of the invention are useful for
XX treating (e.g. gene therapy, protein therapy, vaccine) cancer, such as:
XX lung, colon, colorectal or breast. The present amino acid sequence
XX represents a human cripto protein
XX Sequence 188 AA;
XX Query Match 100.0%; Score 96; DB 5; Length 188;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-06;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 FRDDSIWQPQEEPAIRPR 17
XX 46 FRDDSIWQPQEEPAIRPR 62
XX
XX RESULT 31
XX AAO14636
XX ID AAO14636 standard; protein; 188 AA.
XX AC AAO14636;
XX 13-JUN-2002 (first entry)
XX Human cripto 1 protein.
XX Human; cripto 1 protein; epidermal growth factor family; EGF family;
XX gene therapy; protein therapy; vaccine; lung cancer; colon cancer;
XX colorectal cancer; breast cancer.
XX Homo sapiens.
OS WO200216413-A2.
XX FN
XX PD
XX 28-FEB-2002.
XX 20-AUG-2001; 2001WO-EP009646.
XX 24-AUG-2000; 2000GB-00020953.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX Cassart J, Coche T, Palmantier RM, Vinals Y De BassolsC;
XX WPI; 2002-280910/32.
XX N-PSDB; AAL42196.
XX New Cripto tumor polynucleotides, useful in medicine, particularly for
XX treating tumors (e.g. lung, colorectal, colon or breast tumor) that
XX express a cripto antigen.
XX Claim 1; Page 94; 134pp; English.
XX The invention comprises the nucleotide and amino acid sequences of human
XX cripto proteins. Cripto is a 188 amino acid protein that shares
```

CC homologues with the epidermal growth factor (EGF) family. The crypto  
CC nucleic acid and protein sequences of the invention are useful for  
CC treating (e.g. gene therapy, protein therapy, vaccine) cancer, such as:  
CC lung, colon, colorectal or breast. The present amino acid sequence  
CC represents the human crypto 1 protein  
XX  
SQ Sequence 188 AA;

Query Match 100.0%; Score 96; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWQPEPAIRPR 17  
Db 46 FRDSIWQPEPAIRPR 62  
|||||

RESULT 32  
AAO14727  
ID AAO14727 standard; protein; 188 AA.  
XX  
AC AAO14727;  
XX  
DT 13-JUN-2002 (first entry)  
XX  
DE Human variant crypto 1 protein.  
XX  
KW Human; variant crypto 1 protein; epidermal growth factor family;  
KW EGF family; gene therapy; protein therapy; vaccine; lung cancer;  
KW colon cancer; colorectal cancer; breast cancer; mutant; mutein.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 22 /note= "Wild type Val replaced by Ala"  
XX  
FN WO200216413-A2.  
XX  
PD 28-FEB-2002.  
XX  
XX 20-AUG-2001; 2001WO-EP009646.  
XX  
XX 24-AUG-2000; 2000GB-00020953.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX  
XX Cassart J, Coche T, Palmantier RM, Vinals Y De BassolsC;  
XX  
XX WPI; 2002-280910/32.  
XX N-PSDB; AAL42204.  
XX  
XX New Crypto tumor polynucleotides, useful in medicine, particularly for  
XX treating tumors (e.g. lung, colorectal, colon or breast tumor) that  
XX express a crypto antigen.  
XX  
XX Example 2; Page 105; 134pp; English.  
XX  
XX The invention comprises the nucleotide and amino acid sequences of human  
XX crypto proteins. Crypto is a 188 amino acid protein that shares  
XX homologues with the epidermal growth factor (EGF) family. The crypto  
XX nucleic acid and protein sequences of the invention are useful for  
XX treating (e.g. gene therapy, protein therapy, vaccine) cancer, such as:  
XX lung, colon, colorectal or breast. The present amino acid sequence  
XX represents a variant human crypto 1 protein  
XX  
SQ Sequence 188 AA;

Query Match 100.0%; Score 96; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWQPEPAIRPR 17  
Db 46 FRDSIWQPEPAIRPR 62  
|||||

RESULT 33  
ABP97176  
ID ABP97176 standard; protein; 188 AA.  
XX  
AC ABP97176;  
XX  
DT 01-JUL-2003 (first entry)  
XX  
DE Tumour-associated antigenic target protein TAT177 SEQ ID NO:58.  
XX  
KW Human; tumour-associated antigenic target; TAT; tumour; diagnosis;  
KW cancer.  
XX  
OS Homo sapiens.  
XX  
FN WO2003024392-A2.  
XX  
PD 27-MAR-2003.  
XX  
XX 11-SEP-2002; 2002WO-US028859.  
XX  
XX 18-SEP-2001; 2001US-0323268P.  
XX 19-OCT-2001; 2001US-0339227P.  
XX 07-NOV-2001; 2001US-0336827P.  
XX 20-NOV-2001; 2001US-0331906P.  
XX 02-JAN-2002; 2002US-0345444P.  
XX 03-APR-2002; 2002US-0369724P.  
XX 19-AUG-2002; 2002US-0404809P.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;  
XX Williams PM, Wu TD, Zhang Z;  
XX WPI; 2003-354551/33.  
XX N-PSDB; ACC49494.  
XX  
XX New antibodies against tumor-associated antigenic target polypeptide,  
XX useful for treating or diagnosing tumors or cancers in mammals, e.g.  
XX prostate cancer, lung cancer, prostate adenocarcinomas or renal cell  
XX carcinomas.  
XX  
XX Claim 2; Fig 58; 285pp; English.  
XX  
XX ACC49493 to ACC49552 encode the human tumour-associated antigenic target  
XX (TAT) proteins given in ABP97175 to ABP97234. The present invention  
XX describes an isolated antibody that binds to a polypeptide having at  
XX least 80 % sequence identity to any of the 60 150-800 residue amino acid  
XX sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking  
XX its associated signal peptide, encoded by any of the 60 2000-3000 base  
XX pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have  
XX cytostatic activity. The antibody can be used for treating or diagnosing  
XX tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast  
XX cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal  
XX cell carcinomas, or thyroid cancer  
XX  
SQ Sequence 188 AA;

Query Match 100.0%; Score 96; DB 6; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSIWQPEPAIRPR 17  
Db 46 FRDSIWQPEPAIRPR 62  
|||||

RESULT 34

```
ABP58132
ID ABP58132 standard; protein; 188 AA.
AC ABP58132;
XX
XX 28-MAR-2003 (first entry)
XX Human Cripto CR-3.
XX
XX Cripto; human; antitumour; tumour.
XX
XX Homo sapiens.
XX
XX Location/Qualifiers
XX 31..188
XX Domain /note= "extracellular domain"
XX
XX Region /note= "epitope region specifically described in Claim 8"
XX
XX Domain /note= "ligand/receptor binding domain"
XX
XX Domain /note= "EGF-like domain"
XX
XX Domain /note= "Cysteine-rich domain"
XX
XX Domain /note= "transmembrane domain"
XX
XX WO200288170-A2.
XX
XX 07-NOV-2002.
XX
XX 17-APR-2002; 2002WO-US011950.
XX
XX 26-APR-2001; 2001US-0286782P.
XX 17-MAY-2001; 2001US-0293020P.
XX 26-JUN-2001; 2001US-0301091P.
XX 22-MAR-2002; 2002US-0367002P.
XX
XX (BIOJ ) BIOGEN INC.
XX
XX Sanicola-Nadel M, Williams K, Schiffer S, Rayhorn P;
XX WPI; 2003-120460/11.
XX
XX New Cripto-binding antibodies, useful for modulating the growth of tumor
XX cells, or for treating a subject having e.g. breast, testicular, colon or
XX lung tumor, or suffering from a disorder associated with abnormal levels
XX of Cripto.
XX
XX Claim 2; Page 53; 54pp; English.
XX
XX The present sequence is the amino acid sequence of human Cripto CR-3, a
XX cell surface protein whose over-expression is associated with many tumour
XX types. The invention provides antibodies that specifically bind to Cripto
XX CR-3 or CR-1 (see ABP58131) and which are useful for modulating Cripto
XX signalling or protein interaction, and/or block the interaction between
XX Cripto and ALK4, and/or modulate the growth of tumour cells. The
XX antibodies specifically bind an epitope in the ligand/receptor binding
XX domain of Cripto, especially in an EGF-like domain, or a Cys-rich domain,
XX or a domain spanning amino acid residues 46-62 of Cripto. A claimed
XX composition for administration to a subject having a tumour that
XX expresses Cripto comprises at least one of the antibodies and may also
XX include a chemotherapeutic agent. The composition can be used to modulate
XX the growth of tumour cells in vitro or in vivo, and is used in claimed
XX methods of treating a subject having a tumour which over-expresses Cripto
XX or a condition associated with undesired cell proliferation. The tumour
XX is selected from breast, testicular, colon, lung, ovary, bladder,
XX uterine, cervical, pancreatic and stomach tumour. Also claimed are
XX immunoassay methods of determining whether a tissue expresses Cripto
XX whether a cell line over-expresses Cripto
XX
XX Sequence 188 AA;
```

```
Query Match 100.0%; Score 96; DB 6; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PRDSDIWPQEEPAIRPR 17
DB 46 PRDSDIWPQEEPAIRPR 62
|||||
RESULT 35
ABP58131
ID ABP58131 standard; protein; 188 AA.
XX
XX AC ABP58131;
XX
XX DT 28-MAR-2003 (first entry)
XX
XX DE Human Cripto CR-1.
XX
XX KW Cripto; human; antitumour; tumour.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX Domain 31..188
XX /note= "extracellular domain"
XX
XX FT Region /note= "epitope region specifically described in Claim 8"
XX
XX FT Domain /note= "ligand/receptor binding domain"
XX
XX FT Domain /note= "EGF-like domain"
XX
XX FT Domain /note= "Cysteine-rich domain"
XX
XX FT Domain /note= "transmembrane domain"
XX
XX WO200288170-A2.
XX
XX 07-NOV-2002.
XX
XX 17-APR-2002; 2002WO-US011950.
XX
XX 26-APR-2001; 2001US-0286782P.
XX 17-MAY-2001; 2001US-0293020P.
XX 26-JUN-2001; 2001US-0301091P.
XX 22-MAR-2002; 2002US-0367002P.
XX
XX (BIOJ ) BIOGEN INC.
XX
XX Sanicola-Nadel M, Williams K, Schiffer S, Rayhorn P;
XX WPI; 2003-120460/11.
XX
XX New Cripto-binding antibodies, useful for modulating the growth of tumor
XX cells, or for treating a subject having e.g. breast, testicular, colon or
XX lung tumor, or suffering from a disorder associated with abnormal levels
XX of Cripto.
XX
XX Claim 2; Page 52-53; 54pp; English.
XX
XX The present sequence is the amino acid sequence of human Cripto CR-1, a
XX cell surface protein whose over-expression is associated with many tumour
XX types. The invention provides antibodies that specifically bind to Cripto
XX CR-1 or CR-3 (see ABP58132) and which are useful for modulating Cripto
XX signalling or protein interaction, and/or block the interaction between
XX Cripto and ALK4, and/or modulate the growth of tumour cells. The
XX antibodies specifically bind an epitope in the ligand/receptor binding
XX domain of Cripto, especially in an EGF-like domain, or a Cys-rich domain,
XX or a domain spanning amino acid residues 46-62 of Cripto. A claimed
XX composition for administration to a subject having a tumour that
XX expresses Cripto comprises at least one of the antibodies and may also
XX include a chemotherapeutic agent. The composition can be used to modulate
```



CC the growth of tumour cells in vitro or in vivo, and is used in claimed  
CC methods of treating a subject having a tumour which over-expresses Crip  
CC or a condition associated with undesired cell proliferation. The tumour  
CC is selected from breast, testicular, colon, lung, ovary, bladder,  
CC uterine, cervical, pancreatic and stomach tumour. Also claimed are  
CC immunoassay methods of determining whether a tissue expresses Crip  
CC whether a cell line over-expresses Crip  
XX  
XX

SQ Sequence 188 AA;

Query Match 100.0%; Score 96; DB 6; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | |  
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 36

ADC78799  
ID ADC78799 standard; protein; 188 AA.

XX

AC ADC78799;

XX 01-JAN-2004 (first entry)

XX Human PRO protein #14.

XX human; PRO; cancer; inflammatory bowel disease; ulcerative colitis;  
XX Crohn's disease.

XX Homo sapiens.

XX WO2003034984-A2.

XX 01-MAY-2003.

XX 15-OCT-2002; 2002WO-US033070.

XX 19-OCT-2001; 2001US-0340083P.

XX (GETH ) GENENTECH INC.

XX Goddard A, Gurney AL;

XX WPI; 2003-481990/45.

XX N-PSDB; ADC78798.

XX New PRO polynucleotide and polypeptide, useful for the manufacture of a  
XX medicament for diagnosing or treating cancer or inflammatory bowel  
XX disorder e.g., ulcerative colitis or Crohn's disease.  
XX Claim 12; SEQ ID NO 28; 327pp; English.

XX The invention comprises the amino acid and coding sequences of human PRO

XX proteins. The DNA and protein sequences of the invention are useful for  
XX the diagnosis and treatment of cancer and inflammatory bowel disease  
XX (e.g. ulcerative colitis or Crohn's disease). The present amino acid  
XX sequence represents a human PRO protein of the invention.

SQ Sequence 188 AA;

Query Match 100.0%; Score 96; DB 7; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | |  
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 37

ADC78843

ID ADC78843 standard; protein; 188 AA.

XX

AC ADC78843;

XX 01-JAN-2004 (first entry)

XX Human PRO protein #36.

XX human; PRO; cancer; inflammatory bowel disease; ulcerative colitis;  
XX Crohn's disease.

XX Homo sapiens.

XX WO2003034984-A2.

XX 01-MAY-2003.

XX 15-OCT-2002; 2002WO-US033070.

XX 19-OCT-2001; 2001US-0340083P.

XX (GETH ) GENENTECH INC.

XX Goddard A, Gurney AL;

XX WPI; 2003-481990/45.

XX N-PSDB; ADC78842.

XX New PRO polynucleotide and polypeptide, useful for the manufacture of a  
XX medicament for diagnosing or treating cancer or inflammatory bowel  
XX disorder e.g., ulcerative colitis or Crohn's disease.  
XX Claim 12; SEQ ID NO 72; 327pp; English.

XX The invention comprises the amino acid and coding sequences of human PRO  
XX proteins. The DNA and protein sequences of the invention are useful for  
XX the diagnosis and treatment of cancer and inflammatory bowel disease  
XX (e.g. ulcerative colitis or Crohn's disease). The present amino acid  
XX sequence represents a human PRO protein of the invention.

SQ Sequence 188 AA;

Query Match 100.0%; Score 96; DB 7; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17  
| | | | | | | | | | | | | | | | | | | | |  
Db 46 FRDSDIWPQEEPAIRPR 62

RESULT 38

ADD26059  
ID ADD26059 standard; protein; 188 AA.

XX

AC ADD26059;

XX 15-JAN-2004 (first entry)

XX CR-2 crypto protein.

XX Cripto; Cytostatic; tumour; cell proliferation; antibody.

XX Homo sapiens.

XX WO2003083041-A2.

XX 09-OCT-2003.

XX 01-OCT-2002; 2002WO-US031462.

XX 22-MAR-2002; 2002US-0367002P.

```
PR 24-APR-2002; 2002WO-US011950.
XX (SANI/) SANICOLA-NADEL M.
PA (ADKI/) ADKINS H.
PA (MIKL/) MIKLASZ S D.
PA (RAYH/) RAYHORN P.
PA (SCHI/) SCHIFFER S G.
PA (WILL/) WILLIAMS K.
XX
PI Sanicola-Nadel M, Adkins H, Miklasz SD, Rayhorn P, Schiffer SG;
PI Williams K;
XX
DR WPI; 2003-779452/73.
XX
PT New anti-Cripto antibodies, useful in the therapy of malignant or benign
PT tumors of mammals, or for treating a mammal for a condition associated
PT with undesired cell proliferation.
XX
PS Disclosure; SEQ ID NO 2; 89pp; English.
XX
CC The present invention relates to an antibody that specifically binds to
CC Cripto, to an epitope in the ligand/receptor binding domain of Cripto, or
CC to an epitope comprising the domain spanning amino acid residues 46-62.
CC The antibodies or compositions comprising them are useful for inhibiting
CC growth of tumour cells, treating a mammal having a tumor that
CC overexpresses Cripto, or treating a mammal for a condition associated
CC with undesired cell proliferation. The antibodies are useful in the
CC therapy of malignant or benign tumors of mammals where growth of the
CC tumor is at least partially dependent upon Cripto. The present sequence
CC represents a CR-2 cripto protein of the invention.
XX
SQ Sequence 188 AA;
Query Match 100.0%; Score 96; DB 7; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FRDSDSIWQEEPAIRPR 17
DB 46 FRDSDSIWQEEPAIRPR 62
RESULT 39
ADD26058
ID ADD26058 standard; protein; 188 AA.
XX
AC ADD26058;
XX
DT 15-JAN-2004 (first entry)
XX
DE CR-1 cripto protein.
XX
KW Cripto; Cytostatic; tumour; cell proliferation; antibody.
XX
OS Homo sapiens.
XX
PN WO2003083041-A2.
XX
PD 09-OCT-2003.
XX
PF 01-OCT-2002; 2002WO-US031462.
XX
PR 22-MAR-2002; 2002US-0367002P.
PR 24-APR-2002; 2002WO-US011950.
XX
PA (SANI/) SANICOLA-NADEL M.
PA (ADKI/) ADKINS H.
PA (MIKL/) MIKLASZ S D.
PA (RAYH/) RAYHORN P.
PA (SCHI/) SCHIFFER S G.
PA (WILL/) WILLIAMS K.
XX
PI Sanicola-Nadel M, Adkins H, Miklasz SD, Rayhorn P, Schiffer SG;
PI Williams K;
XX
DR WPI; 2003-779452/73.
XX
PT New anti-Cripto antibodies, useful in the therapy of malignant or benign
PT tumors of mammals, or for treating a mammal for a condition associated
PT with undesired cell proliferation.
XX
PS Disclosure; SEQ ID NO 2; 89pp; English.
XX
CC The present invention relates to an antibody that specifically binds to
CC Cripto, to an epitope in the ligand/receptor binding domain of Cripto, or
CC to an epitope comprising the domain spanning amino acid residues 46-62.
CC The antibodies or compositions comprising them are useful for inhibiting
CC growth of tumour cells, treating a mammal having a tumor that
CC overexpresses Cripto, or treating a mammal for a condition associated
CC with undesired cell proliferation. The antibodies are useful in the
CC therapy of malignant or benign tumors of mammals where growth of the
CC tumor is at least partially dependent upon Cripto. The present sequence
CC represents a CR-2 cripto protein of the invention.
XX
SQ Sequence 188 AA;
Query Match 100.0%; Score 96; DB 7; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FRDSDSIWQEEPAIRPR 17
DB 46 FRDSDSIWQEEPAIRPR 62
RESULT 40
ADI82176
ID ADI82176 standard; protein; 188 AA.
XX
AC ADI82176;
XX
DT 22-APR-2004 (first entry)
XX
DE Human teratocarcinoma-derived growth factor 2.
XX
KW Human; embryonic stem cell; pluripotent stem cell; abnormal cell growth;
KW malignancy; differentiation.
XX
OS Homo sapiens.
XX
PN US2003224411-A1.
XX
PD 04-DEC-2003.
XX
PF 13-MAR-2003; 2003US-00388578.
XX
PR 13-MAR-2003; 2003US-00388578.
XX
PA (STAN/) STANTON L W.
PA (BRAN/) BRANDENBERGER R.
PA (GOLD/) GOLD J D.
PA (IRVI/) IRVING J M.
PA (MAND/) MANDALAM R.
PA (MOKM/) MOK M.
PA (SHEL/) SHELTON D.
XX
PI Stanton LW, Brandenberger R, Gold JD, Irving JM, Mandalam R;
PI Mok M, Shelton D;
XX
DR WPI; 2004-119701/12.
DR N-PSDB; ADI82175.
XX
PT Assessing culture of undifferentiated primate pluripotent stem cells by
PT detecting expression of markers e.g., Zic family member 3, other than
PT human telomerase reverse transcriptase/octamer binding transcription
PT factor.
```

XX Claim 1; SEQ ID NO 6; 106pp; English.

XX The invention relates to assessing a culture of undifferentiated primate

CC pluripotent stem cells (pPS, e.g. embryonic stem cells), involving

CC detecting expression of markers (MR1) e.g. Zic family member 3 (ZIC3), as

CC given in specification, other than human telomerase reverse transcriptase

CC (hTERT) or octamer binding transcription factor (Oct3/4, or a marker

CC (MR2) such as Cripto or podocalyxin-like protein and hTERT and/or Oct3/4

CC or second marker chosen from (MR2). Also included are maintaining (M2)

CC pPS cells in a pluripotent state (involves causing them to express one of

CC the following markers (MR3) at a higher level, FOXO1A, ZIC3, hypochelical

CC protein FLJ20582, Forkhead box H1 (FOXH1), Zinc finger protein, Hsdl2,

CC KRAB-zinc finger protein SZFI-1 or zinc finger protein of cerebellum

CC ZIC2, or any other marker (MR4) chosen from PRD protein Jade-1 (Jade-1),

CC kruppel-like zinc finger protein (ZNF300), etc., as given in the

CC specification), causing pPS cells to differentiate into a particular

CC tissue type by causing them to express one of the markers chosen from

CC (MR3) or (MR4) or markers chosen from GATA binding protein 3 (GATA3),

CC core promoter element binding protein (COPEB), etc., as given in the

CC specification), maintaining pPS cells in a pluripotent state (involves

CC culturing pPS cells or their progeny in the presence of a normally

CC secreted protein that is encoded by a gene that down-regulated upon

CC differentiation of human embryonic stem (hES) cells, chosen from

CC Fibrillin 3 gene, LEFT B gene, ZIC3 gene, EPHA1 gene, etc., as given in

CC the specification), causing pPS cells to differentiate (involves

CC culturing pPS cells or their progeny in the presence of a normally

CC secreted protein that is encoded by a gene that up-regulated upon

CC differentiation of hES cells, chosen from p311 protein gene, Tax

CC interaction protein 1 gene, KIA0853 protein gene, keratin 19 (KRT 19)

CC gene, etc., as given in the specification), causing an encoding sequence

CC to be preferentially expressed in undifferentiated pPS cells, causing an

CC encoding sequence to be preferentially expressed in differentiated cells,

CC sorting (M4) differentiated cells from less differentiated cells

CC (involves separating cells expressing a surface marker chosen from any

CC one of MR1 from cells not expressing the marker), causing pPS cells to

CC proliferate without differentiation, identifying genes that are up or

CC down regulated during differentiation of pPS cells, and a kit (I) for

CC assessing a culture of pPS cells by M1. The method (M1) is useful for

CC assessing culture of undifferentiated primate pluripotent stem cells and

CC for assessing the growth characteristics of a cell population. The cell

CC population has been obtained by culturing cells from human blastocyst or

CC from a human patient suspected of having a clinical condition related to

CC abnormal cell growth. The method further involves determining whether the

CC cell population is pluripotent from the marker expression and assessing

CC whether the patient has a malignancy from the marker expression. The

CC present sequence is a protein whose expression is down regulated in

CC pluripotent stem cells.

XX

SQ Sequence 188 AA;

Query Match 100.0%; Score 96; DB 8; Length 188;

Best Local Similarity 100.0%; Pred. No. 1.6e-06;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSDIWQOEPAIRPR 17

Db 46 FRDSDIWQOEPAIRPR 62

RESULT 41

ID ADR70486

XX ADR70486 standard; protein; 188 AA.

XX

AC ADR70486;

XX

XX

DT 02-DEC-2004 (first entry)

XX

DE Human teratocarcinoma-derived growth factor (Cripto).

XX

XX Human; telomerase reverse transcriptase; TERT;

XX POU domain, class 5 transcription factor; POU5F1; Oct3; Oct4;

XX teratocarcinoma-derived growth factor; Cripto; podocalyxin-like; PODXL;

KW gastrin-releasing peptide receptor; GRPR; human embryonic stem cell; hES;

KW primate pluripotent stem cell; cancer; gene expression; cell separation;

XX differentiation.

OS Homo sapiens.

XX

XX US2004180347-A1.

PN

XX

PD 16-SEP-2004.

XX

PF 13-MAR-2003; 2003US-00389431.

XX

PR 13-MAR-2003; 2003US-00389431.

XX

XX (STAN// STANTON L W.

PA (BRAN// BRANDENBERGER R.

PA (GOLD// GOLD J D.

PA (TRVI// IRVING J M.

PA (MAND// MANDALAM R.

PA (MOKM// MOK M.

XX

XX Stanton LW, Brandenberger R, Gold JD, Irving JM, Mandalam R;

PI Mok M;

XX

DR WPI; 2004-675599/66.

XX N-PSDB; ADR70485.

DR

XX

XX Assessing culture of undifferentiated human embryonic stem cells or their

PT progeny, by detecting Cripto, gastrin-releasing peptide (GRP) receptor

PT and podocalyxin-like protein markers, and either hTERT and/or Oct3/4, or

PT GRP receptor.

PT

XX

XX Disclosure; SEQ ID NO 6; 57pp; English.

XX

XX The invention relates to assessing a culture of undifferentiated human

CC embryonic stem (hES) cells (undifferentiated primate pluripotent stem

CC cells) or their progeny, involves detecting or measuring a marker such as

CC Cripto (teratocarcinoma-derived growth factor), gastrin-releasing peptide

CC (GRP) receptor and podocalyxin-like protein, and either hTERT (telomerase

CC reverse transcriptase) and/or Oct3/4 (also known as POU domain, class 5,

CC transcription factor 1(POU5F1)), or GRP receptor. The method involves

CC detecting or measuring at least two markers, and detecting or measuring

CC hTERT and/or Oct3/4. The expression of the marker(s) is detected or

CC measured at mRNA level by PCR amplification. The expression of the

CC marker(s) is detected or measured at the protein level by antibody assay.

CC The method involves quantifying the proportion of undifferentiated hES

CC cells or differentiated cells in the culture from the marker expression.

CC The level of the marker is determined to be at least 100-fold higher than

CC the level of the marker in BJ fibroblasts or is determined to be no less

CC than 100-fold lower than the level of the marker in hES cells, cultured

CC on an extracellular matrix in medium conditioned with mouse embryonic

CC fibroblasts and containing 4 ng/ml basic fibroblast growth factor. The

CC method further involves modifying the culture conditions so as to cause

CC the hES cells to increase expression of the marker detected or measured

CC in the culture. The method is useful for assessing a culture of

CC undifferentiated hES cells or their progeny. The marker used in the above

CC method is useful for characterising pluripotent stem cells and their

CC differentiated progeny, for clinical diagnosis of cancer, for assessing

CC and manipulating culture conditions, regulating gene expression, cell

CC separation and purification, and to influence differentiation. The

CC present sequence is a marker protein of the invention for

CC undifferentiated stem cells.

XX

SQ Sequence 188 AA;

Query Match 100.0%; Score 96; DB 8; Length 188;

Best Local Similarity 100.0%; Pred. No. 1.6e-06;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRDSDIWQOEPAIRPR 17

Db 46 FRDSDIWQOEPAIRPR 62

RESULT 42  
ADS88697  
ID ADS88697 standard; protein; 188 AA.  
XX AC  
XX ADT79198;  
XX DT 30-DEC-2004 (first entry)  
XX AC  
XX ADT79198;  
XX DT 16-DEC-2004 (first entry)  
XX AC  
XX ADT79198;  
XX DT Amino acid sequence of a human Cripto protein.  
XX AC  
XX ADT79198;  
XX DT stem cell differentiation; cardiomyocyte; stem cell;  
XX AC epidermal growth factor; EGF; EGF-CFC family; Cripto; neuronal cell;  
XX ADT79198; Alq4; Nodal; heart disease; neuropathology; myocardial infarction;  
XX DT Parkinson's disease; Alzheimer's disease; retinal degeneration;  
XX AC cell therapy.  
XX DT Homo sapiens.  
XX AC OS  
XX ADT79198; WO2004083375-A2.  
XX PN  
XX ADT79198; 30-SEP-2004.  
XX PD  
XX ADT79198; 19-MAR-2004; 2004WO-IT000133.  
XX PF  
XX ADT79198; 21-MAR-2003; 2003IT-RM000125.  
XX PR 29-JUL-2003; 2003IT-RM000370.  
XX AC  
XX ADT79198; (MINC/) MINCHIOTTI G.  
XX PA (PERS/) PERSICO M.  
XX ADT79198; (PARI/) PARISI S.  
XX PA  
XX ADT79198; Minchiotti G, Persico M, Parisi S;  
XX PI  
XX ADT79198; WPI; 2004-691035/67.  
XX DR  
XX ADT79198; Inducing stem cell differentiation in cardiomyocytes or neuronal cells by  
XX PT exposure to a protein of EGF-CFC or Cripto protein inhibitors,  
XX PT respectively, useful in treating heart diseases and neuropathologies.  
XX PT  
XX PT Disclosure; Page 13; 35pp; English.  
XX PS  
XX ADT79198; The specification describes a method for inducing stem cell  
XX CC differentiation into cardiomyocytes. The method comprises exposing the  
XX CC cells to a protein of the epidermal growth factor (EGF)-CFC family or its  
XX CC derivatives and having at least the EGF and CFC domains (e.g. Cripto  
XX CC protein). A second method is described for stem cell differentiation into  
XX CC neuronal cells, where the cells are exposed to an inhibitor of the Cripto  
XX CC protein. The Cripto protein inhibitor is an anti-Cripto antibody or its  
XX CC functional fragments, or is a peptide specifically selected from a random  
XX CC combinatorial peptide library, or is an antagonist of the Alq4(receptor)-  
XX CC Cripto(co-receptor)-Nodal(ligand) pathway. The differentiated stem cells  
XX CC of the invention are useful in the treatment of heart diseases and  
XX CC neuropathologies, e.g. myocardial infarction, Parkinson's disease,  
XX CC Alzheimer's disease and retinal degeneration. They are especially useful  
XX CC for cell therapy. The Cripto protein or its inhibitors is useful in the  
XX CC preparation of a composition able to direct stem cell differentiation  
XX CC toward the neuronal lineage. The present sequence represents a Cripto  
XX CC protein, which may be used in the method of the invention.  
XX SQ  
XX Sequence 188 AA;  
XX Query Match 100.0%; Score 96; DB 8; Length 188;  
XX Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 FRDSSIWPOEPAIRPR 17  
Db 46 FRDSSIWPOEPAIRPR 62  
RESULT 43  
ADT79198

ID ADT79198 standard; protein; 188 AA.  
XX AC  
XX ADT79198;  
XX DT 30-DEC-2004 (first entry)  
XX AC  
XX ADT79198;  
XX DE Human Cripto 3 protein.  
XX KW Cripto protein; vaccine; cancer; lung; colon; colorectal; breast; T cell;  
XX KW immune response; cytostatic; human.  
XX OS Homo sapiens.  
XX AC US2004202648-A1.  
XX PN 14-OCT-2004.  
XX PD  
XX ADT79198; 01-APR-2004; 2004US-00816476.  
XX PF  
XX ADT79198; 04-APR-2003; 2003US-0460733P.  
XX PR  
XX ADT79198; (CABE/) CABEZON T B V S.  
XX PA (GERA/) GERARD C M G.  
XX ADT79198; (PALM/) PALMANTIER R M.  
XX PA (VINA/) VINALS Y D B C.  
XX ADT79198; Cabazon TEVS, Gerard CMG, Palmantier RM, Vinals YDBC;  
XX PI  
XX ADT79198; WPI; 2004-727978/71.  
XX DR N-PSDB; ADT79198.  
XX ADT79198; Novel immunogenic fragment of Cripto polypeptide, useful as vaccine for  
XX PT treating cancer e.g. colon, lung, colorectal and breast cancer.  
XX PT  
XX PT Claim 1; SEQ ID NO 4; 64pp; English.  
XX PS  
XX ADT79198; The invention relates to immunogenic fragments of Cripto polypeptide  
XX CC which are used as vaccines. The immunogenic fragment of Cripto protein is  
XX CC useful for treating or diagnosing cancer preferably lung, colon,  
XX CC colorectal or breast cancer, for stimulating T cells specific for Cripto  
XX CC and for inhibiting the development of a cancer in a patient. It is useful  
XX CC for inducing an immune response to Cripto in an animal. The present  
XX CC sequence is the human Cripto 3 protein.  
XX SQ  
XX Sequence 188 AA;  
XX Query Match 100.0%; Score 96; DB 8; Length 188;  
XX Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 FRDSSIWPOEPAIRPR 17  
Db 46 FRDSSIWPOEPAIRPR 62  
RESULT 44  
ADT79200  
ID ADT79200 standard; protein; 188 AA.  
XX AC  
XX ADT79200;  
XX AC  
XX ADT79200; 30-DEC-2004 (first entry)  
XX DT  
XX ADT79200; Human Cripto protein.  
XX DE  
XX ADT79200; Cripto protein; vaccine; cancer; lung; colon; colorectal; breast; T cell;  
XX KW immune response; cytostatic; human; gene; ds; variant; mutant.  
XX KW  
XX OS Homo sapiens.  
XX AC US2004202648-A1.  
XX PN 14-OCT-2004.  
XX PD

```
XX 01-APR-2004; 2004US-00816476.
XX 04-APR-2003; 2003US-0460733P.
XX (CABE/) CABEZON T E V S.
XX (GERA/) GERARD C M G.
XX (PALM/) PALMANTIER R M.
XX (VINA/) VINALS Y D B C.
XX Cabezon TEVS, Gerard CMG, Palmantier RM, Vinals YDBC;
XX WPI; 2004-727978/71.
XX N-PSDB; ADT79199.
XX Novel immunogenic fragment of Cripto polypeptide, useful as vaccine for
XX treating cancer e.g. colon, lung, colorectal and breast cancer.
XX Example 2; SEQ ID NO 6; 64pp; English.
XX The invention relates to immunogenic fragments of Cripto polypeptide
XX which are used as vaccines. The immunogenic fragment of Cripto protein is
XX useful for treating or diagnosing cancer preferably lung, colon,
XX colorectal or breast cancer, for stimulating T cells specific for Cripto
XX and for inhibiting the development of a cancer in a patient. It is useful
XX for inducing an immune response to Cripto in an animal. The present
XX sequence is the human Cripto protein.
XX Sequence 188 AA;
XX Query Match 100.0%; Score 96; DB 8; Length 188;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-06;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 FRDSSIWPQEEPAIRPR 17
XX Db 46 FRDSSIWPQEEPAIRPR 62
XX
XX RESULT 45
XX ADT79197
XX ID ADT79197 standard; protein; 188 AA.
XX AC ADT79197;
XX XX 30-DEC-2004 (first entry)
XX DE Human Cripto 1 protein.
XX KW Cripto protein; vaccine; cancer; lung; colon; colorectal; breast; T cell;
XX KW immune response; cytostatic; human.
XX OS Homo sapiens.
XX XX US2004202648-A1.
XX PN 14-OCT-2004.
XX PD 01-APR-2004; 2004US-00816476.
XX PF 04-APR-2003; 2003US-0460733P.
XX PR (CABE/) CABEZON T E V S.
XX PA (GERA/) GERARD C M G.
XX PA (PALM/) PALMANTIER R M.
XX PA (VINA/) VINALS Y D B C.
XX PI Cabezon TEVS, Gerard CMG, Palmantier RM, Vinals YDBC;
XX XX WPI; 2004-727978/71.
XX DR N-PSDB; ADT79289.
XX XX Novel immunogenic fragment of Cripto polypeptide, useful as vaccine for
XX treating cancer e.g. Colon, lung, colorectal and breast cancer.
XX Example 2; SEQ ID NO 96; 64pp; English.
XX The invention relates to immunogenic fragments of Cripto polypeptide
XX which are used as vaccines. The immunogenic fragment of Cripto protein is
XX useful for treating or diagnosing cancer preferably lung, colon,
XX colorectal or breast cancer, for stimulating T cells specific for Cripto
XX and for inhibiting the development of a cancer in a patient. It is useful
XX for inducing an immune response to Cripto in an animal. The present
XX sequence is the human Cripto 1 variant protein.
XX Sequence 188 AA;
XX Query Match 100.0%; Score 96; DB 8; Length 188;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-06;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 FRDSSIWPQEEPAIRPR 17
XX Db 46 FRDSSIWPQEEPAIRPR 62
XX
XX RESULT 45
XX ADT79197
XX ID ADT79197 standard; protein; 188 AA.
XX AC ADT79197;
XX XX 30-DEC-2004 (first entry)
XX DE Human Cripto 1 protein.
XX KW Cripto protein; vaccine; cancer; lung; colon; colorectal; breast; T cell;
XX KW immune response; cytostatic; human.
XX OS Homo sapiens.
XX XX US2004202648-A1.
XX PN 14-OCT-2004.
XX PD 01-APR-2004; 2004US-00816476.
XX PF 04-APR-2003; 2003US-0460733P.
XX PR (CABE/) CABEZON T E V S.
XX PA (GERA/) GERARD C M G.
XX PA (PALM/) PALMANTIER R M.
XX PA (VINA/) VINALS Y D B C.
XX PI Cabezon TEVS, Gerard CMG, Palmantier RM, Vinals YDBC;
XX XX WPI; 2004-727978/71.
XX DR N-PSDB; ADT79195.
XX XX Novel immunogenic fragment of Cripto polypeptide, useful as vaccine for
```

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSSIWPQEPAIRPR 17  
|||||  
Db 46 FRDSSIWPQEPAIRPR 62

RESULT 47  
ADY85961  
ID ADY85961 standard; protein; 188 AA.

XX AC ADY85961;

XX 16-JUN-2005 (first entry)

XX Human Cripto oncoprotein Seq 2.

XX cell signaling; oncogenesis; antisense therapy; cytostatic; Cripto;  
XX mutagenesis; protein interaction; oncoprotein.

XX Homo sapiens.

XX Key Location/Qualifiers  
XX Modified-site 88

XX /note= "Modified by O-fucosylation"

XX WO2005028433-A2.

XX 31-MAR-2005.

XX 14-SEP-2004; 2004WO-US029967.

XX 15-SEP-2003; 2003US-0503046P.

XX (RERE-) RES DEV FOUND.

XX Vale W, Gray PC, Harrison CA;

XX WPI; 2005-242562/25.

XX Augmenting signaling of a ligand of receptor serine kinase in a cell  
XX comprises inhibiting the formation of complexes between Cripto and the  
XX ligand on the surface of the cell.

XX Disclosure; SEQ ID NO 2; 60pp; English.

XX This invention relates to a novel method for augmenting signaling of a  
XX ligand of a receptor serine kinase within a cell by inhibiting the  
XX formation of complexes between Cripto and this ligand on the surface of  
XX the cell. Specifically, it refers to TGF-beta and activin which are the  
XX ligands of serine kinase receptors and which regulate tissue homeostasis  
XX by activating the Smad2/3 intracellular signaling pathway; disruption of  
XX this signaling pathway is associated with oncogenesis and tumorigenesis.  
XX As such, the present invention describes a method for augmenting Smad2/3  
XX signaling in a cell by administering a mutant ligand that retains  
XX antagonism by Cripto. Note that augmentation of signaling increases  
XX phosphorylation and activation of Smad2 and Smad3 in the cell, such that  
XX it decreases the proliferative rate of the cell. The receptor serine  
XX kinase is a type I activin receptor-like kinases-4 or -5 (ALK-4 or ALK-5)  
XX and the formation of complexes is inhibited by suppressing expression of  
XX Cripto using antisense oligonucleotides (siRNA) directed against Cripto,  
XX and also mutating at least one allele of Cripto by homologous  
XX recombination. Accordingly, pharmaceutical compositions derived thereof  
XX exhibit cytostatic activity. This polypeptide sequence is the human  
XX Cripto oncoprotein of the invention.

XX Sequence 188 AA;

Query Match 100.0%; Score 96; DB 9; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSSIWPQEPAIRPR 17  
|||||  
Db 46 FRDSSIWPQEPAIRPR 62

RESULT 48  
ADZ42238

ID ADZ42238 standard; protein; 188 AA.

XX AC ADZ42238;

XX 30-JUN-2005 (first entry)

XX Human Cripto-1 protein.

XX neurodegenerative disease; neuroprotective; gene therapy.

XX Homo sapiens.

XX WO2005033341-A2.

XX 14-APR-2005.

XX 01-OCT-2004; 2004WO-US032649.

XX 03-OCT-2003; 2003US-0508750P.

XX (UNIV ) UNIV KANSAS MEDICAL CENT.

XX (USSH ) US DEPT OF HEALTH & HUMAN SERVICES.

XX Salomon D, Berman N, Stephens E;

XX WPI; 2005-273509/28.

XX N-PSDB; ADZ42237.

XX Detecting a neurodegenerative disease, such as NeuroAIDS, Alzheimer's  
XX disease, multiple sclerosis, Parkinson's disease and encephalitis, by  
XX assaying the copy number of a Cripto-1 gene or the expression level of a  
XX Cripto-1 gene product.

XX Disclosure; SEQ ID NO 2; 42pp; English.

XX The invention relates to a novel method for detecting a neurodegenerative  
XX disease in a mammal. The method comprises assaying the copy number of a  
XX Cripto-1 gene (also known as Teratocarcinoma-derived Growth Factor-1  
XX (TDGF-1) gene) or the expression level of a Cripto-1 gene product in the  
XX central nervous system of the mammal, where an amplification of the  
XX Cripto-1 gene or an overexpression of the Cripto-1 gene product is  
XX indicative of a neurodegenerative disease in the mammal. The invention  
XX further comprises: a method for inhibiting progression of a  
XX neurodegenerative disease in a mammal, comprising administering to the  
XX mammal an agent that inhibits Cripto-1 in the central nervous system of  
XX the mammal, where the progression of the neurodegenerative disease is  
XX inhibited; and an isolated or purified oligonucleotide consisting  
XX essentially of SEQ ID NO: 3 (ADZ42239) or 4 (ADZ42240). The agent may be  
XX used in the treatment of neurodegenerative disorders by gene therapy. The  
XX methods and compositions of the present invention are useful for the  
XX diagnosis, prevention and/or treatment of neurodegenerative disorders,  
XX such as NeuroAIDS, Alzheimer's disease, multiple sclerosis, amyotrophic  
XX lateral sclerosis (ALS), Parkinson's disease, and encephalitis. This  
XX sequence represents the human Cripto-1 protein of the invention.

XX Sequence 188 AA;

Query Match 100.0%; Score 96; DB 9; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSSIWPQEPAIRPR 17  
|||||  
Db 46 FRDSSIWPQEPAIRPR 62

```
RESULT 49
ABB77114
ID ABB77114 standard; protein; 360 AA.
AC
XX
AC ABB77114;
XX
DT 08-OCT-2002 (first entry)
XX
XX Human Cripto-1/IgG1 Fc domain fusion protein mutant CR38(cc)-FcT88A.
DE
XX
XX Human; Cripto-1; CR-1; mutant; tumour; fusion protein; IgG1; Fc domain;
KW cell proliferation.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX WO200222808-A2.
XX
XX 21-MAR-2002.
XX
XX 18-SEP-2001; 2001WO-US029066.
XX
XX 18-SEP-2000; 2000US-0233148P.
XX
XX (BIOJ ) BIOGEN INC.
XX
XX Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;
XX WPI; 2002-339868/37.
XX
XX New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful
XX for treating cell proliferation, especially cancer, comprises amino acid
XX change that prevents fucosylation at Thr88.
XX
XX Disclosure; Page 36; 41pp; English.
XX
XX The sequence represents a C-terminal truncated form of human Cripto-1
XX fused to the Fc domain of human IgG1. The invention relates to a novel
XX mutant CRIPTO polypeptide, or its functional fragment, having at least
XX one amino acid alteration at positions 86, 87 or 88. The mutant
XX polypeptide, or its chimera, is used to inhibit growth of tumour cells,
XX in vivo or in vitro, particularly for treating breast, ovarian, renal,
XX colorectal, uterine, prostatic, lung, bladder or central nervous system
XX cancers, melanoma and leukaemia, also generally for treating undesired
XX cell proliferation
XX
XX Sequence 360 AA;

Query Match 100.0%; Score 96; DB 5; Length 360;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17
Db 9 FRDSDIWPQEEPAIRPR 25

Search completed: February 3, 2006, 18:50:57
Job time : 354 secs

cell proliferation.
KW
XX
XX Homo sapiens.
OS Synthetic.
XX
XX WO200222808-A2.
XX
XX 21-MAR-2002.
XX
XX 18-SEP-2001; 2001WO-US029066.
XX
XX 18-SEP-2000; 2000US-0233148P.
XX
XX (BIOJ ) BIOGEN INC.
XX
XX Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;
XX WPI; 2002-339868/37.
XX
XX New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful
XX for treating cell proliferation, especially cancer, comprises amino acid
XX change that prevents fucosylation at Thr88.
XX
XX Disclosure; Page 36; 41pp; English.
XX
XX The sequence represents a C-terminal truncated form of human Cripto-1
XX fused to the Fc domain of human IgG1. The invention relates to a novel
XX mutant CRIPTO polypeptide, or its functional fragment, having at least
XX one amino acid alteration at positions 86, 87 or 88. The mutant
XX polypeptide, or its chimera, is used to inhibit growth of tumour cells,
XX in vivo or in vitro, particularly for treating breast, ovarian, renal,
XX colorectal, uterine, prostatic, lung, bladder or central nervous system
XX cancers, melanoma and leukaemia, also generally for treating undesired
XX cell proliferation
XX
XX Sequence 360 AA;

Query Match 100.0%; Score 96; DB 5; Length 360;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17
Db 9 FRDSDIWPQEEPAIRPR 25

Search completed: February 3, 2006, 18:50:57
Job time : 354 secs

cell proliferation.
KW
XX
XX Homo sapiens.
OS Synthetic.
XX
XX WO200222808-A2.
XX
XX 21-MAR-2002.
XX
XX 18-SEP-2001; 2001WO-US029066.
XX
XX 18-SEP-2000; 2000US-0233148P.
XX
XX (BIOJ ) BIOGEN INC.
XX
XX Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;
XX WPI; 2002-339868/37.
XX
XX New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful
XX for treating cell proliferation, especially cancer, comprises amino acid
XX change that prevents fucosylation at Thr88.
XX
XX Disclosure; Page 36; 41pp; English.
XX
XX The sequence represents a mutant C-terminal truncated form of human
XX Cripto-1 fused to the Fc domain of human IgG1, where the threonine
XX residue at position 51 has been replaced with an alanine residue. The
XX invention relates to a novel mutant CRIPTO polypeptide, or its functional
XX fragment, having at least one amino acid alteration at positions 86, 87
XX or 88. The mutant polypeptide, or its chimera, is used to inhibit growth
XX of tumour cells, in vivo or in vitro, particularly for treating breast,
XX ovarian, renal, colorectal, uterine, prostatic, lung, bladder or central
XX nervous system cancers, melanoma and leukaemia, also generally for
XX treating undesired cell proliferation
XX
XX Sequence 360 AA;

Query Match 100.0%; Score 96; DB 5; Length 360;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRDSDIWPQEEPAIRPR 17
Db 9 FRDSDIWPQEEPAIRPR 25

RESULT 50
ABB77107
ID ABB77107 standard; protein; 360 AA.
XX
XX
XX ABB77107;
XX
DT 08-OCT-2002 (first entry)
XX
XX Human Cripto-1/IgG1 Fc domain fusion protein (CR38(cc)-Fc).
DE
XX
XX Human; Cripto-1; CR-1; mutant; tumour; fusion protein; IgG1; Fc domain;
```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:30:21 ; Search time 45.9556 Seconds  
(without alignments)  
338.219 Million cell updates/sec

Title: US-10-693-538-1  
Perfect score: 1053  
Sequence: 1 MDCRKMARSYSYVIMWMAIS.....SARTTFMLVGLICLSQSY 188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5 COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/6 COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/H COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/RE COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	188	1	US-07-749-001-3
2	1053	100.0	188	1	US-08-154-198-3
3	1053	100.0	188	1	US-08-463-335-3
4	1053	100.0	188	1	US-08-464-023A-3
5	1053	100.0	188	1	US-08-471-371-7
6	1049	99.6	192	2	US-09-949-016-8959
7	1049	99.6	192	2	US-09-949-016-11113
8	1021	97.0	187	6	5256643-2
9	1013	96.2	188	1	US-07-749-001-5
10	1013	96.2	188	1	US-08-154-198-5
11	1013	96.2	188	1	US-08-463-335-5
12	1013	96.2	188	1	US-08-464-023A-5
13	1013	96.2	192	2	US-09-949-016-7062
14	955	90.7	174	6	5256643-3
15	629	59.7	160	1	US-08-441-629-7
16	629	59.7	160	2	US-08-776-207-7
17	629	59.7	160	2	US-09-507-773-7
18	629	59.7	160	2	US-10-016-447-7
19	629	59.7	160	4	PCT-US95-09172-7
20	547	51.9	94	1	US-08-208-008C-14
21	371	35.2	68	1	US-08-208-008C-13
22	274.5	26.1	190	1	US-08-441-629-4
23	274.5	26.1	190	2	US-08-776-207-4
24	274.5	26.1	190	2	US-09-507-773-4
25	274.5	26.1	190	2	US-10-016-447-4
26	274.5	26.1	190	4	PCT-US95-09172-4
27	233	22.1	230	1	US-08-471-371-2

28 135 12.8 26 1 US-08-208-008C-12 Sequence 12, Appli  
29 129 12.3 379 2 US-09-907-794A-4 Sequence 4, Appli  
30 129 12.3 379 2 US-09-905-125A-4 Sequence 4, Appli  
31 129 12.3 379 2 US-09-902-775A-4 Sequence 4, Appli  
32 129 12.3 379 2 US-09-906-700-4 Sequence 4, Appli  
33 129 12.3 379 2 US-09-903-603A-4 Sequence 4, Appli  
34 129 12.3 379 2 US-09-904-920A-4 Sequence 4, Appli  
35 129 12.3 379 2 US-09-909-064-4 Sequence 4, Appli  
36 129 12.3 379 2 US-09-905-381A-4 Sequence 4, Appli  
37 129 12.3 379 2 US-09-906-618-4 Sequence 4, Appli  
38 129 12.3 379 2 US-09-906-646-4 Sequence 4, Appli  
39 129 12.3 379 2 US-09-904-462-4 Sequence 4, Appli  
40 129 12.3 379 2 US-09-902-736A-4 Sequence 4, Appli  
41 129 12.3 379 2 US-09-906-722A-4 Sequence 4, Appli  
42 124 11.8 380 2 US-09-205-258-441 Sequence 441, App  
43 124 11.8 380 2 US-10-004-860-441 Patent No. 5256643  
44 123 11.7 20 6 5256643-6  
45 119 11.3 2556 1 US-08-185-432-17 Sequence 17, Appli

ALIGNMENTS

RESULT 1  
US-07-749-001-3  
; Sequence 3, Application US/07749001  
; Patent No. 5264557  
; GENERAL INFORMATION:  
; APPLICANT: Salomon, David S.  
; APPLICANT: Persico, Maria G.  
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
; NUMBER OF INVENTIONS: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/749,001  
; FILING DATE: 19910823  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Scott, Watson T.  
; REGISTRATION NUMBER: 26,581  
; REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 248453 CUSH  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 188 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-749-001-3  
Query Match 100.0%; Score 1053; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 66-96;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDCRKMARSYSYVIMWMAISKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
Db 1 MDCRKMARSYSYVIMWMAISKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
Qy 61 PRSSQRPVPMGIQHSKELNRTCLNGTCTMLGSCACPPSPFYGRNCEHDKVRCNGSVPH 120

Db 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCACPPSFYGRNCEHDVKEKCGSVPH 120  
Qy 121 DTWLPKCKSLCKCKWHGQRCFPQAFPLPGCDGLVNDHLVASRTPELPPSARTTTFMLVGI 180  
Db 121 DTWLPKCKSLCKCKWHGQRCFPQAFPLPGCDGLVNDHLVASRTPELPPSARTTTFMLVGI 180  
Qy 181 CLSIQSY 188  
Db 181 CLSIQSY 188

RESULT 2  
US-08-154-198-3  
; Sequence 3, Application US/08154198  
; Patent No. 5620866  
; GENERAL INFORMATION:  
; APPLICANT: SALOMON, David S.  
; APPLICANT: PERSICO, Maria G.  
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/154,198  
; FILING DATE: 17-NOV-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,001  
; FILING DATE: 23-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 15280-63-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 188 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-154-198-3

Query Match 100.0%; Score 1053; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6e-96;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDCRKMARFSYSVTWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEPAIR 60  
Db 1 MDCRKMARFSYSVTWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEPAIR 60  
Qy 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCACPPSFYGRNCEHDVKEKCGSVPH 120  
Db 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCACPPSFYGRNCEHDVKEKCGSVPH 120  
Qy 121 DTWLPKCKSLCKCKWHGQRCFPQAFPLPGCDGLVNDHLVASRTPELPPSARTTTFMLVGI 180  
Db 121 DTWLPKCKSLCKCKWHGQRCFPQAFPLPGCDGLVNDHLVASRTPELPPSARTTTFMLVGI 180  
Qy 181 CLSIQSY 188

Db 181 CLSIQSY 188  
RESULT 3  
US-08-463-335-3  
; Sequence 3, Application US/08463335  
; Patent No. 5650285  
; GENERAL INFORMATION:  
; APPLICANT: Salomon, David S.  
; APPLICANT: Persico, Maria G.  
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DABBY & CUSHMAN  
; STREET: 1615 L Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,335  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,001  
; FILING DATE: 23-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Scott, Watson T.  
; REGISTRATION NUMBER: 26,581  
; REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 248453 CUSH  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 188 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-463-335-3  
Query Match 100.0%; Score 1053; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6e-96;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDCRKMARFSYSVTWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEPAIR 60  
Db 1 MDCRKMARFSYSVTWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEPAIR 60  
Qy 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCACPPSFYGRNCEHDVKEKCGSVPH 120  
Db 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCACPPSFYGRNCEHDVKEKCGSVPH 120  
Qy 121 DTWLPKCKSLCKCKWHGQRCFPQAFPLPGCDGLVNDHLVASRTPELPPSARTTTFMLVGI 180  
Db 121 DTWLPKCKSLCKCKWHGQRCFPQAFPLPGCDGLVNDHLVASRTPELPPSARTTTFMLVGI 180  
Qy 181 CLSIQSY 188  
Db 181 CLSIQSY 188  
RESULT 4  
US-08-464-023A-3  
; Sequence 3, Application US/08464023A  
; Patent No. 5854399

GENERAL INFORMATION:  
APPLICANT: Salomon, David S.  
TITLE OF INVENTION: A Human Crypto-Related Gene  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Steuart Street Tower  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105-1492  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,023A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,001  
FILING DATE: 23-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/154,198  
FILING DATE: 17-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 15280-63-1-1  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-464-023A-3

Query Match 100.0%; Score 1053; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6e-96;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDCRKMARFSYSVIWIMAIKVPFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60  
Db 1 MDCRKMARFSYSVIWIMAIKVPFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60

Qy 61 PRSSQRPVPMGIIQHSKELNRTCCINGTCLMGSCFACPPSYGRNCEHVDVRKENCSSVPH 120  
Db 61 PRSSQRPVPMGIIQHSKELNRTCCINGTCLMGSCFACPPSYGRNCEHVDVRKENCSSVPH 120

Qy 121 DTWLPKKSICKWHGQLRCFPAFLPGCDGLVNDHLVASRTPPELPPSARTTTFMVLGI 180  
Db 121 DTWLPKKSICKWHGQLRCFPAFLPGCDGLVNDHLVASRTPPELPPSARTTTFMVLGI 180

Qy 181 CLSIQSYI 188  
Db 181 CLSIQSYI 188

RESULT 5  
US-08-471-371-7  
Sequence 7, Application US/08471371  
Patent No. 5981215  
GENERAL INFORMATION:  
APPLICANT: Meissner, Paul S.  
TITLE OF INVENTION: HUMAN CRIPTIN GROWTH FACTOR  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CRECHT,  
ADDRESSEE: STEWART & OLSTEIN  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: N.J.  
COUNTRY: U.S.A  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,371  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Periato, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 32580-455  
TELEPHONE: (201) 994-1700  
TELEFAX: (201) 994-1744  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 188 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-471-371-7

Query Match 100.0%; Score 1053; DB 1; Length 188;  
Best Local Similarity 100.0%; Pred. No. 6e-96;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDCRKMARFSYSVIWIMAIKVPFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60  
Db 1 MDCRKMARFSYSVIWIMAIKVPFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60

Qy 61 PRSSQRPVPMGIIQHSKELNRTCCINGTCLMGSCFACPPSYGRNCEHVDVRKENCSSVPH 120  
Db 61 PRSSQRPVPMGIIQHSKELNRTCCINGTCLMGSCFACPPSYGRNCEHVDVRKENCSSVPH 120

Qy 121 DTWLPKKSICKWHGQLRCFPAFLPGCDGLVNDHLVASRTPPELPPSARTTTFMVLGI 180  
Db 121 DTWLPKKSICKWHGQLRCFPAFLPGCDGLVNDHLVASRTPPELPPSARTTTFMVLGI 180

Qy 181 CLSIQSYI 188  
Db 181 CLSIQSYI 188

RESULT 6  
US-09-949-016-8959  
Sequence 8959, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8959

```
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8959

Query Match      99.6%; Score 1049; DB 2; Length 192;
Best Local Similarity 99.5%; Pred. No. 1.5e-95;
Matches 187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  1 MDCRKMARFSYSVTIWIWMAISKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEEPAIR 60
Db  5 MDCRKMARFSYSVTIWIWMAISKAFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEEPAIR 64
Qy  61 PRSSORVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEDHVRKENGCSVPH 120
Db  65 PRSSORVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEDHVRKENGCSVPH 124
Qy  121 DTWLPKKCSLCKCKWHGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180
Db  125 DTWLPKKCSLCKCKWHGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 184
Qy  181 CLSIQSY 188
Db  185 CLSIQSY 192

RESULT 7
US-09-949-016-11113
; Sequence 11113, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11113
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11113

Query Match      99.6%; Score 1049; DB 2; Length 192;
Best Local Similarity 99.5%; Pred. No. 1.5e-95;
Matches 187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  1 MDCRKMARFSYSVTIWIWMAISKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEEPAIR 60
Db  5 MDCRKMARFSYSVTIWIWMAISKAFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEEPAIR 64
Qy  61 PRSSORVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEDHVRKENGCSVPH 120
Db  65 PRSSORVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEDHVRKENGCSVPH 124
Qy  121 DTWLPKKCSLCKCKWHGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180
Db  125 DTWLPKKCSLCKCKWHGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 184
Qy  181 CLSIQSY 188
Db  185 CLSIQSY 192

RESULT 8
```

```
5256643--2
; Patent No. 5256643
; APPLICANT: Persico, Maria G.; Salomon, David S.
; TITLE OF INVENTION: HUMAN CRYPTO PROTEIN
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/530,165
; FILING DATE: 29-MAY-1990
; SEQ ID NO:2
; LENGTH: 187
5256643--2

Query Match      97.0%; Score 1021; DB 6; Length 187;
Best Local Similarity 98.4%; Pred. No. 8.3e-93;
Matches 184; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy  1 MDCRKMARFSYSVTIWIWMAISKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEEPAIR 60
Db  1 MDCRKMARFSYSVTIWIWMAISKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEEPAIR 60
Qy  61 PRSSORVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEDHVRKENGCSVPH 120
Db  61 PRSSORVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEDHVRKENGCSVPH 120
Qy  121 DTWLPKKCSLCKCKWHGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180
Db  121 DTWLPKKCSLCKCKWHGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180
Qy  181 CLSIQSY 187
Db  181 CLSIQSY 187

RESULT 9
US-07-749-001-5
; Sequence 5, Application US/07749001
; Patent No. 5264557
; GENERAL INFORMATION:
; APPLICANT: Salomon, David S.
; APPLICANT: Persico, Maria G.
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/749,001
; FILING DATE: 19910823
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 248453 CUSH
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-749-001-5
```

```
Query Match          96.2%; Score 1013; DB 1; Length 188;
Best Local Similarity 96.8%; Pred. No. 5.1e-92;
Matches 182; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MDCRKMRFYSYVITWIMAIKSKVPELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60
DB 1 MDCRKMRFYSYVITWIMAIKSKVPELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60

QY 61 PRSQRVPPMGIOHSHKELNRTCCINGGTCLMGSCFACPPSFYGRNCEHVDVRKNCGSVPH 120
DB 61 PRSQRVPPMGIOHSHKELNRTCCINGGTCLMGSCFACPPSFYGRNCEHVDVRKNCGSVPH 120

QY 121 DTWLPKKCSICKWHGQRCFPOAFLPGCDGLVMDHLVASRTPPELPPSARTTTFMLVGI 180
DB 121 DTWLPKKCSICKWHGQRCFPOAFLPGCDGLVMDHLVASRTPPELPPSARTTTFMLVGI 180

QY 181 CLSIQSY 188
DB 181 CLSIQSY 188

RESULT 10
US-08-154-198-5
; Sequence 5, Application US/08154198
; Patent No. 5620866
; GENERAL INFORMATION:
; APPLICANT: SALOMON, David S.
; APPLICANT: PERSICO, Maria G.
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 17-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,001
; FILING DATE: 23-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 15280-63-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-154-198-5

Query Match          96.2%; Score 1013; DB 1; Length 188;
Best Local Similarity 96.8%; Pred. No. 5.1e-92;
Matches 182; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MDCRKMRFYSYVITWIMAIKSKVPELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60
DB 1 MDCRKMRFYSYVITWIMAIKSKVPELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60

QY 61 PRSQRVPPMGIOHSHKELNRTCCINGGTCLMGSCFACPPSFYGRNCEHVDVRKNCGSVPH 120
DB 61 PRSQRVPPMGIOHSHKELNRTCCINGGTCLMGSCFACPPSFYGRNCEHVDVRKNCGSVPH 120

QY 121 DTWLPKKCSICKWHGQRCFPOAFLPGCDGLVMDHLVASRTPPELPPSARTTTFMLVGI 180
DB 121 DTWLPKKCSICKWHGQRCFPOAFLPGCDGLVMDHLVASRTPPELPPSARTTTFMLVGI 180

QY 181 CLSIQSY 188
DB 181 CLSIQSY 188

RESULT 11
US-08-463-335-5
; Sequence 5, Application US/08463335
; Patent No. 5650285
; GENERAL INFORMATION:
; APPLICANT: Salomon, David S.
; APPLICANT: Persico, Maria G.
; TITLE OF INVENTION: A HUMAN CRYPTO-RELATED GENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,001
; FILING DATE: 23-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/91630/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 248453 CUSH
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-335-5

Query Match          96.2%; Score 1013; DB 1; Length 188;
Best Local Similarity 96.8%; Pred. No. 5.1e-92;
Matches 182; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MDCRKMRFYSYVITWIMAIKSKVPELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60
DB 1 MDCRKMRFYSYVITWIMAIKSKVPELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60

QY 61 PRSQRVPPMGIOHSHKELNRTCCINGGTCLMGSCFACPPSFYGRNCEHVDVRKNCGSVPH 120
DB 61 PRSQRVPPMGIOHSHKELNRTCCINGGTCLMGSCFACPPSFYGRNCEHVDVRKNCGSVPH 120

QY 121 DTWLPKKCSICKWHGQRCFPOAFLPGCDGLVMDHLVASRTPPELPPSARTTTFMLVGI 180
DB 121 DTWLPKKCSICKWHGQRCFPOAFLPGCDGLVMDHLVASRTPPELPPSARTTTFMLVGI 180

QY 181 CLSIQSY 188
DB 181 CLSIQSY 188
```

Qy 181 CLSIQSY 188  
 Db 181 CLSIQSY 188

RESULT 12

US-08-464-023A-5  
 ; Sequence 5, Application US/08464023A  
 ; Patent No. 5854399  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Salomon, David S.  
 ; TITLE OF INVENTION: A Human Crypto-Related Gene  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend and Crew  
 ; STREET: One Market Plaza, Steuart Street Tower  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94105-1492  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/464,023A  
 ; FILING DATE: 05-JUN-1995  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/749,001  
 ; FILING DATE: 23-AUG-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/154,198  
 ; FILING DATE: 17-NOV-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bastian, Kevin L.  
 ; REGISTRATION NUMBER: 34,774  
 ; REFERENCE/DOCKET NUMBER: 15280-63-1-1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 543-9600  
 ; TELEFAX: (415) 543-5043  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 188 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-464-023A-5

Query Match 96.2%; Score 1013; DB 1; Length 188;  
 Best Local Similarity 96.8%; Pred. No. 5.1e-92;  
 Matches 182; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MDCRQARFYSYVIWMAISKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPQEEPAIR 60  
 Db 1 MDCRQARFYSYVIWMAISKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPQEEPAIR 60  
 Qy 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVVRKENGCSVPH 120  
 Db 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVVRKENGCSVPH 120  
 Qy 121 DTWLPKCKSLCKCWGHQQLRCFPQAFPGCDGLVMDHLVASRTPELPPSARTTTTFLMVG 180  
 Db 121 DTWLPKCKSLCKCWGHQQLRCFPQAFPGCDGLVMDHLVASRTPELPPSARTTTTFLMVG 180  
 Qy 181 CLSIQSY 188  
 Db 181 CLSIQSY 188

RESULT 13

US-09-949-016-7062  
 ; Sequence 7062, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7062  
 ; LENGTH: 192  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-09-949-016-7062

Query Match 96.2%; Score 1013; DB 2; Length 192;  
 Best Local Similarity 96.8%; Pred. No. 5.3e-92;  
 Matches 182; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MDCRQARFYSYVIWMAISKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPQEEPAIR 60  
 Db 5 MDCRQARFYSYVIWMAISKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPQEEPAIR 64  
 Qy 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVVRKENGCSVPH 120  
 Db 65 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVVRKENGCSVPH 124  
 Qy 121 DTWLPKCKSLCKCWGHQQLRCFPQAFPGCDGLVMDHLVASRTPELPPSARTTTTFLMVG 180  
 Db 125 DTWLPKCKSLCKCWGHQQLRCFPQAFPGCDGLVMDHLVASRTPELPPSARTTTTFLMVG 184  
 Qy 181 CLSIQSY 188  
 Db 185 CLSIQSY 192

RESULT 14  
 5256643-3  
 ; Patent No. 5256643  
 ; APPLICANT: Persico, Maria G.; Salomon, David S.  
 ; TITLE OF INVENTION: HUMAN CRYPTO PROTEIN  
 ; NUMBER OF SEQUENCES: 18  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/530,165  
 ; FILING DATE: 29-MAY-1990  
 ; SEQ ID NO:3:  
 ; LENGTH: 174  
 5256643-3

Query Match 90.7%; Score 955; DB 6; Length 174;  
 Best Local Similarity 98.8%; Pred. No. 2.4e-86;  
 Matches 169; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 18 AISKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPQEEPAIRPRSSQRPVPMGIQHSKE 77  
 Db 4 AISKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPQEEPAIRPRSSQRPVPMGIQHSKE 63  
 Qy 78 LNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVVRKENGCSVPHDTWLPKCKSLCKCWGHQ 137  
 Db 64 LNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVVRKENGCSVPHDTWLPKCKSLCKCWGHQ 123  
 Qy 138 LRCPPQAFPGCDGLVMDHLVASRTPELPPSARTTTTFLMVGICLSIQSY 188  
 Db 124 LRCPPQAFPGCDGLVMDHLVASRTPELPPSARTTTTFLMVGICLSIQSY 174

RESULT 15  
US-08-441-629-7  
; Sequence 7, Application US/08441629  
; Patent No. 5766923  
; GENERAL INFORMATION:  
; APPLICANT: Kirschner, Marc W.  
; APPLICANT: Kinoshita, No. 5766923iyuki  
; TITLE OF INVENTION: RECEPTOR-LIGAND ASSAY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/441,629  
; FILING DATE: 15-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/279,217  
; FILING DATE: 22-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: HU95-01A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 861-6240  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 160 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-441-629-7

Query Match 59.7%; Score 629; DB 1; Length 160;  
Best Local Similarity 68.0%; Pred. No. 2.6e-54;  
Matches 117; Conservative 14; Mismatches 29; Indels 12; Gaps 2;  
  
QY 6 MARFSYSVIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEPAIRPRSSQ 65  
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
1 MGYFSSVLLVAISSAFEGPVAGRD-----LAIRDNSIWDQKEPAVRDRSTQ 49  
  
QY 66 RVPFMGIQHSKELNRITCCLNGGTCLGSCFACPPSFYGRNCEHDVRKENCOSVPHDTWLP 125  
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
50 FVPSVGIONSLSNKTCTCLNGGTCLGSCFACPPSFYGRNCEHDVRKENCOSVPHDTWLP 109  
  
QY 126 KCSLCKCWGQRCFQALPGCDGLVMDHLVASRTPELPSPARTTTFML 177  
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
110 KCSLRCRWGQLHCLPQTFLPGCDGHVMDODLKASRTPCQTPSV-TTTFML 160

Search completed: February 3, 2006, 18:33:12  
Job time : 46.9556 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:31:31 ; Search time 146.222 Seconds  
(without alignments)  
537.209 Million cell updates/sec

Title: US-10-693-538-1  
Perfect score: 1053  
Sequence: 1 MDCRKRARFSYVITWIMATIS.....SARTTTFMLVIGICLSIQSY 188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main:  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pap:\*  
2: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pap:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pap:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pap:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pap:\*  
6: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	188	4	US-10-241-220-58
2	1053	100.0	188	4	US-10-388-578-6
3	1053	100.0	188	4	US-10-390-566-1
4	1053	100.0	188	4	US-10-362-597A-3
5	1053	100.0	188	4	US-10-362-597A-6
6	1053	100.0	188	4	US-10-407-481-3
7	1053	100.0	188	4	US-10-407-481-6
8	1053	100.0	188	4	US-10-712-124-58
9	1053	100.0	188	4	US-10-693-538-1
10	1053	100.0	188	4	US-10-389-431-6
11	1053	100.0	188	4	US-10-816-476-3
12	1053	100.0	188	4	US-10-816-476-6
13	1053	100.0	188	5	US-10-872-372-58
14	1053	100.0	188	5	US-10-872-991-58
15	1053	100.0	188	5	US-10-491-997-28
16	1053	100.0	188	5	US-10-940-431-2
17	1053	100.0	188	5	US-10-983-340-13
18	1049	99.6	188	4	US-10-362-597A-96
19	1049	99.6	188	4	US-10-407-481-96
20	1049	99.6	188	4	US-10-816-476-96
21	1048	99.5	188	4	US-10-390-566-9
22	1013	96.2	188	4	US-10-390-566-2
23	1013	96.2	188	4	US-10-362-597A-4
24	1013	96.2	188	4	US-10-407-481-4
25	1013	96.2	188	4	US-10-693-538-2
26	1013	96.2	188	4	US-10-816-476-4
27	1013	96.2	188	5	US-10-491-997-72

Sequence 22, Appl  
Sequence 7, Appl  
Sequence 66, Appl  
Sequence 6, Appl  
Sequence 3, Appl  
Sequence 13, Appl  
Sequence 10, Appl  
Sequence 19, Appl  
Sequence 16, Appl  
Sequence 23, Appl  
Sequence 26, Appl  
Sequence 7, Appl  
Sequence 14, Appl  
Sequence 4, Appl  
Sequence 11, Appl  
Sequence 20, Appl  
Sequence 27, Appl  
Sequence 17, Appl

ALIGNMENTS

RESULT 1  
US-10-241-220-58  
; Sequence 58, Application US/10241220  
; Publication No. US20030148408A1  
; GENERAL INFORMATION:  
; APPLICANT: Frantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P5010R1-US  
; CURRENT APPLICATION NUMBER: US/10/241,220  
; CURRENT FILING DATE: 2002-12-13  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 58  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-241-220-58

Query Match 100.0%; Score 1053; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.2e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDCRKRARFSYVITWIMATISKVFELGLVAGLGHQEFARPSRGYLAFRDDSIMPQEEPAIR 60  
DB 1 MDCRKRARFSYVITWIMATISKVFELGLVAGLGHQEFARPSRGYLAFRDDSIMPQEEPAIR 60  
QY 61 PRSSQRVPVPMGIQHSKELNRTCLNGTCLMGSCFACPPSFYGRNCEHDVKNKCGSVPH 120  
DB 61 PRSSQRVPVPMGIQHSKELNRTCLNGTCLMGSCFACPPSFYGRNCEHDVKNKCGSVPH 120  
QY 121 DTWLPPKCSLCKCWGHQQLRCFQOAFPGCDGLVMDHLVASRTPPELPSPSARTTTFMLVGI 180  
DB 121 DTWLPPKCSLCKCWGHQQLRCFQOAFPGCDGLVMDHLVASRTPPELPSPSARTTTFMLVGI 180  
QY 181 CLSIQSY 188  
DB 181 CLSIQSY 188

RESULT 2  
US-10-388-578-6  
; Sequence 6, Application US/10388578  
; Publication No. US20030224411A1

```

; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Stanton, Lawrence
; APPLICANT: Ralph, Brandenberger
; APPLICANT: Joseph, Gold D.
; APPLICANT: John, Irving
; APPLICANT: Mandalam, Ramkumar
; APPLICANT: Mok, Michael
; APPLICANT: Shelton, Dawne
; TITLE OF INVENTION: Genes that are Up- or Down-Regulated During Differentiation of Hu
; TITLE OF INVENTION: Embryonic Stem Cells
; FILE REFERENCE: 135/001
; CURRENT APPLICATION NUMBER: US/10/388.578
; CURRENT FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: Custom
; SEQ ID NO 6
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-388-578-6

Query Match      100.0%; Score 1053; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.2e-93;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDCRKFARFSYVIMWAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60
Db 1 MDCRKFARFSYVIMWAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60

Qy 61 PRSSQRPVPMGIQHSKELNRTCCNGGTCMLGSCFACPPSYGRNCEHDVRKENGCSVPH 120
Db 61 PRSSQRPVPMGIQHSKELNRTCCNGGTCMLGSCFACPPSYGRNCEHDVRKENGCSVPH 120

Qy 121 DTWLPKCKSLCKWHGQLRCFPPQAFPLGCDGLVMDHLVASRTPELPPSARTTTFMLVGI 180
Db 121 DTWLPKCKSLCKWHGQLRCFPPQAFPLGCDGLVMDHLVASRTPELPPSARTTTFMLVGI 180

Qy 181 CLSIQSY 188
Db 181 CLSIQSY 188

RESULT 3
US-10-390-566-1
; Sequence 1, Application US/10390566
; Publication No. US20030232755A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin P.
; APPLICANT: Foley, Susan
; APPLICANT: Schiffer, Susan
; APPLICANT: Domon, Bruno
; APPLICANT: Sanicola-Nadel, Michele
; TITLE OF INVENTION: Crypto Mutant and Uses Thereof
; FILE REFERENCE: BGN-A104
; CURRENT APPLICATION NUMBER: US/10/390,566
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/233,148
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US01/29066
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-390-566-1

Query Match      100.0%; Score 1053; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.2e-93;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDCRKFARFSYVIMWAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60
Db 1 MDCRKFARFSYVIMWAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60

Qy 61 PRSSQRPVPMGIQHSKELNRTCCNGGTCMLGSCFACPPSYGRNCEHDVRKENGCSVPH 120
Db 61 PRSSQRPVPMGIQHSKELNRTCCNGGTCMLGSCFACPPSYGRNCEHDVRKENGCSVPH 120

Qy 121 DTWLPKCKSLCKWHGQLRCFPPQAFPLGCDGLVMDHLVASRTPELPPSARTTTFMLVGI 180
Db 121 DTWLPKCKSLCKWHGQLRCFPPQAFPLGCDGLVMDHLVASRTPELPPSARTTTFMLVGI 180

Qy 181 CLSIQSY 188
Db 181 CLSIQSY 188

RESULT 4
US-10-362-597A-3
; Sequence 3, Application US/10362597A
; Publication No. US20040054142A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237
; CURRENT APPLICATION NUMBER: US/10/362,597A
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: PCT/EP01/0964
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Human
US-10-362-597A-3

Query Match      100.0%; Score 1053; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.2e-93;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDCRKFARFSYVIMWAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60
Db 1 MDCRKFARFSYVIMWAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60

Qy 61 PRSSQRPVPMGIQHSKELNRTCCNGGTCMLGSCFACPPSYGRNCEHDVRKENGCSVPH 120
Db 61 PRSSQRPVPMGIQHSKELNRTCCNGGTCMLGSCFACPPSYGRNCEHDVRKENGCSVPH 120

Qy 121 DTWLPKCKSLCKWHGQLRCFPPQAFPLGCDGLVMDHLVASRTPELPPSARTTTFMLVGI 180
Db 121 DTWLPKCKSLCKWHGQLRCFPPQAFPLGCDGLVMDHLVASRTPELPPSARTTTFMLVGI 180

Qy 181 CLSIQSY 188
Db 181 CLSIQSY 188

RESULT 5
US-10-362-597A-6
; Sequence 6, Application US/10362597A
; Publication No. US20040054142A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237
; CURRENT APPLICATION NUMBER: US/10/362,597A
```

```
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: PCT/EP01/0964
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 188
; TYPE: PRT
; ORGANISM: human
US-10-362-597A-6

Query Match      100.0%; Score 1053; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.2e-93;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCRKARFYSYVIWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWQEEPAIR 60
   |||||||
Db 1 MDCRKARFYSYVIWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWQEEPAIR 60
   |||||||

QY 61 PRSSQRPVPMGIOHSKELNRTCCNGTCLMGSCFACPPSFYGRNCEHDVVRKENCSSVPH 120
   |||||||
Db 61 PRSSQRPVPMGIOHSKELNRTCCNGTCLMGSCFACPPSFYGRNCEHDVVRKENCSSVPH 120
   |||||||

QY 121 DTWLPKKCSLCKWHGQLRCFQAFPGCDGLVWDEHLVASRTPPELPPSARTTTFMLVGI 180
   |||||||
Db 121 DTWLPKKCSLCKWHGQLRCFQAFPGCDGLVWDEHLVASRTPPELPPSARTTTFMLVGI 180
   |||||||

QY 181 CLSIQSY 188
   |||||||
Db 181 CLSIQSY 188
   |||||||

RESULT 6
US-10-407-481-3
; Sequence 3, Application US/10407481
; Publication No. US20040138112A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237-1
; CURRENT APPLICATION NUMBER: US/10/407,481
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: PCT/EP01/09646
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Human
US-10-407-481-3

Query Match      100.0%; Score 1053; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.2e-93;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCRKARFYSYVIWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWQEEPAIR 60
   |||||||
Db 1 MDCRKARFYSYVIWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWQEEPAIR 60
   |||||||

QY 61 PRSSQRPVPMGIOHSKELNRTCCNGTCLMGSCFACPPSFYGRNCEHDVVRKENCSSVPH 120
   |||||||
Db 61 PRSSQRPVPMGIOHSKELNRTCCNGTCLMGSCFACPPSFYGRNCEHDVVRKENCSSVPH 120
   |||||||

QY 121 DTWLPKKCSLCKWHGQLRCFQAFPGCDGLVWDEHLVASRTPPELPPSARTTTFMLVGI 180
   |||||||
Db 121 DTWLPKKCSLCKWHGQLRCFQAFPGCDGLVWDEHLVASRTPPELPPSARTTTFMLVGI 180
   |||||||

QY 181 CLSIQSY 188
   |||||||
Db 181 CLSIQSY 188
   |||||||

RESULT 7
US-10-407-481-6
; Sequence 6, Application US/10407481
; Publication No. US20040138112A1
; GENERAL INFORMATION:
; APPLICANT: CASSART, Jean-Pol
; APPLICANT: COCHE, Thierry
; APPLICANT: PALMANTER, Remi M.
; APPLICANT: VINALS Y DE BASSOLS, Carlota
; TITLE OF INVENTION: Vaccines
; FILE REFERENCE: B45237-1
; CURRENT APPLICATION NUMBER: US/10/407,481
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: PCT/EP01/09646
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 188
; TYPE: PRT
; ORGANISM: human
US-10-407-481-6

Query Match      100.0%; Score 1053; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.2e-93;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCRKARFYSYVIWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWQEEPAIR 60
   |||||||
Db 1 MDCRKARFYSYVIWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWQEEPAIR 60
   |||||||

QY 61 PRSSQRPVPMGIOHSKELNRTCCNGTCLMGSCFACPPSFYGRNCEHDVVRKENCSSVPH 120
   |||||||
Db 61 PRSSQRPVPMGIOHSKELNRTCCNGTCLMGSCFACPPSFYGRNCEHDVVRKENCSSVPH 120
   |||||||

QY 121 DTWLPKKCSLCKWHGQLRCFQAFPGCDGLVWDEHLVASRTPPELPPSARTTTFMLVGI 180
   |||||||
Db 121 DTWLPKKCSLCKWHGQLRCFQAFPGCDGLVWDEHLVASRTPPELPPSARTTTFMLVGI 180
   |||||||

QY 181 CLSIQSY 188
   |||||||
Db 181 CLSIQSY 188
   |||||||

RESULT 8
US-10-712-124-58
; Sequence 58, Application US/10712124
; Publication No. US20040146907A1
; GENERAL INFORMATION:
; APPLICANT: SMITH, VICTORIA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING DYSPLASIA
; FILE REFERENCE: P2000R1
; CURRENT APPLICATION NUMBER: US/10/712,124
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 60/425,813
; PRIOR FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 123
; SEQ ID NO 58
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-712-124-58

Query Match      100.0%; Score 1053; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.2e-93;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCRKARFYSYVIWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWQEEPAIR 60
   |||||||
Db 1 MDCRKARFYSYVIWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWQEEPAIR 60
   |||||||
```

Qy 61 PRSSQVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEHDVREKNCGSVPH 120  
Db 61 PRSSQVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEHDVREKNCGSVPH 120  
Qy 121 DTWLPKCSLCKCWGHQQLRCFPQAFPLPGCDGLVMDHLVASRTPELPPSARTTTFMVLGI 180  
Db 121 DTWLPKCSLCKCWGHQQLRCFPQAFPLPGCDGLVMDHLVASRTPELPPSARTTTFMVLGI 180  
Qy 181 CLSIQSY 188  
Db 181 CLSIQSY 188

## RESULT 9

US-10-693-538-1  
; Sequence 1, Application US/10693538  
; Publication No. US20040146940A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc.  
; APPLICANT: Sanicola-Nadel, Michele  
; APPLICANT: Williams, Kevin  
; APPLICANT: Schiffer, Susan  
; APPLICANT: Rayhorn, Paul  
; TITLE OF INVENTION: Cripito Blocking Antibodies and Uses  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: BIN-A117CN  
; CURRENT APPLICATION NUMBER: US/10/693,538  
; PRIOR FILING DATE: 2003-10-23  
; PRIOR APPLICATION NUMBER: PCT/US02/11950  
; PRIOR FILING DATE: 2002-04-16  
; PRIOR APPLICATION NUMBER: 60/286,782  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: 60/293,020  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/301,091  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/367,002  
; PRIOR FILING DATE: 2002-03-22  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-693-538-1

Query Match 100.0%; Score 1053; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.2e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDCRMRFSYSVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDSDIWPQEEPAIR 60  
Db 1 MDCRMRFSYSVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDSDIWPQEEPAIR 60  
Qy 61 PRSSQVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEHDVREKNCGSVPH 120  
Db 61 PRSSQVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEHDVREKNCGSVPH 120  
Qy 121 DTWLPKCSLCKCWGHQQLRCFPQAFPLPGCDGLVMDHLVASRTPELPPSARTTTFMVLGI 180  
Db 121 DTWLPKCSLCKCWGHQQLRCFPQAFPLPGCDGLVMDHLVASRTPELPPSARTTTFMVLGI 180  
Qy 181 CLSIQSY 188  
Db 181 CLSIQSY 188

## RESULT 10

US-10-389-431-6  
; Sequence 6, Application US/10389431  
; Publication No. US20040180347A1  
; GENERAL INFORMATION:

; APPLICANT: Geron Corporation  
; APPLICANT: Stanton, Lawrence  
; APPLICANT: Ralph, Brandenberger  
; APPLICANT: Joseph, Gold D.  
; APPLICANT: John, Irving  
; APPLICANT: Mandalam, Ramkumar  
; APPLICANT: Mok, Michael  
; TITLE OF INVENTION: A Marker System for Preparing and Characterizing High-Quality Human Embryonic Stem Cells  
; TITLE OF INVENTION: Embryonic Stem Cells  
; FILE REFERENCE: 135/002  
; CURRENT APPLICATION NUMBER: US/10/389,431  
; CURRENT FILING DATE: 2003-03-13  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-389-431-6

Query Match 100.0%; Score 1053; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.2e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDCRMRFSYSVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDSDIWPQEEPAIR 60  
Db 1 MDCRMRFSYSVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDSDIWPQEEPAIR 60  
Qy 61 PRSSQVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEHDVREKNCGSVPH 120  
Db 61 PRSSQVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEHDVREKNCGSVPH 120  
Qy 121 DTWLPKCSLCKCWGHQQLRCFPQAFPLPGCDGLVMDHLVASRTPELPPSARTTTFMVLGI 180  
Db 121 DTWLPKCSLCKCWGHQQLRCFPQAFPLPGCDGLVMDHLVASRTPELPPSARTTTFMVLGI 180  
Qy 181 CLSIQSY 188  
Db 181 CLSIQSY 188

## RESULT 11

US-10-816-476-3  
; Sequence 3, Application US/10816476  
; Publication No. US20040202648A1  
; GENERAL INFORMATION:  
; APPLICANT: CABEZON, Teresa Elisa Virginia Silva  
; APPLICANT: GERARD, Catherine Marie Ghislaine  
; APPLICANT: PALMANTIER, Remi M.  
; APPLICANT: VINALS Y DE BASSOLS, Carlota  
; TITLE OF INVENTION: Vaccines  
; FILE REFERENCE: VU60111  
; CURRENT APPLICATION NUMBER: US/10/816,476  
; CURRENT FILING DATE: 2004-04-01  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE: 2004-04-01  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Human  
US-10-816-476-3

Query Match 100.0%; Score 1053; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.2e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDCRMRFSYSVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDSDIWPQEEPAIR 60  
Db 1 MDCRMRFSYSVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDSDIWPQEEPAIR 60  
Qy 61 PRSSQVPPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEHDVREKNCGSVPH 120

Db 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCSDHVRKENCSSVPH 120  
Qy 121 DTWLPKCSLCKWHGQRCFPPQAFPLGCDGLVNDHVLVASRTPELPPSARTTTTFLVLGI 180  
Db 121 DTWLPKCSLCKWHGQRCFPPQAFPLGCDGLVNDHVLVASRTPELPPSARTTTTFLVLGI 180  
Qy 181 CLSIQSY 188  
Db 181 CLSIQSY 188

RESULT 12

US-10-816-476-6  
; Sequence 6, Application US/10816476  
; Publication No. US20040202648A1  
; GENERAL INFORMATION:  
; APPLICANT: CABEZON, Teresa Elisa Virginia Silva  
; APPLICANT: GERARD, Catherine Marie Ghislaine  
; APPLICANT: PALMANTIER, Remi M.  
; APPLICANT: VINALS Y DE BASSOLS, Carlota  
; TITLE OF INVENTION: Vaccines  
; FILE REFERENCE: VU60111  
; CURRENT APPLICATION NUMBER: US/10/816,476  
; PRIOR FILING DATE: 2004-04-01  
; PRIOR APPLICATION NUMBER:  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: human  
US-10-816-476-6

Query Match 100.0%; Score 1053; DB 4; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.2e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDCRKMARFYSYVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
Db 1 MDCRKMARFYSYVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
Qy 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCSDHVRKENCSSVPH 120  
Db 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCSDHVRKENCSSVPH 120  
Qy 121 DTWLPKCSLCKWHGQRCFPPQAFPLGCDGLVNDHVLVASRTPELPPSARTTTTFLVLGI 180  
Db 121 DTWLPKCSLCKWHGQRCFPPQAFPLGCDGLVNDHVLVASRTPELPPSARTTTTFLVLGI 180  
Qy 181 CLSIQSY 188  
Db 181 CLSIQSY 188

RESULT 13

US-10-872-972-58  
; Sequence 58, Application US/10872972  
; Publication No. US20040229277A1  
; GENERAL INFORMATION:  
; APPLICANT: Frantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P5010R1-US  
; CURRENT APPLICATION NUMBER: US/10/872,972

; CURRENT FILING DATE: 2004-06-21  
; PRIOR APPLICATION NUMBER: US/10/241,220  
; PRIOR FILING DATE: 2002-09-11  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 58  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-872-972-58

Query Match 100.0%; Score 1053; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.2e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDCRKMARFYSYVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
Db 1 MDCRKMARFYSYVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
Qy 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCSDHVRKENCSSVPH 120  
Db 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCSDHVRKENCSSVPH 120  
Qy 121 DTWLPKCSLCKWHGQRCFPPQAFPLGCDGLVNDHVLVASRTPELPPSARTTTTFLVLGI 180  
Db 121 DTWLPKCSLCKWHGQRCFPPQAFPLGCDGLVNDHVLVASRTPELPPSARTTTTFLVLGI 180  
Qy 181 CLSIQSY 188  
Db 181 CLSIQSY 188

RESULT 14

US-10-872-991-58  
; Sequence 58, Application US/10872991  
; Publication No. US20040242860A1  
; GENERAL INFORMATION:  
; APPLICANT: Frantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P5010R1-US  
; CURRENT APPLICATION NUMBER: US/10/872,991  
; CURRENT FILING DATE: 2004-06-21  
; PRIOR APPLICATION NUMBER: US/10/241,220  
; PRIOR FILING DATE: 2002-09-11  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 58  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-872-991-58

Query Match 100.0%; Score 1053; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.2e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDCRKMARFYSYVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
Db 1 MDCRKMARFYSYVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
Qy 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCSDHVRKENCSSVPH 120  
Db 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCSDHVRKENCSSVPH 120  
Qy 121 DTWLPKCSLCKWHGQRCFPPQAFPLGCDGLVNDHVLVASRTPELPPSARTTTTFLVLGI 180  
Db 121 DTWLPKCSLCKWHGQRCFPPQAFPLGCDGLVNDHVLVASRTPELPPSARTTTTFLVLGI 180

Qy 181 CLSIQSY 188  
| | | | |  
Db 181 CLSIQSY 188

## RESULT 15

US-10-491-997-28  
; Sequence 28, Application US/10491997  
; Publication No. US20050089957A1  
; GENERAL INFORMATION:  
; APPLICANT: Gurney, Audrey  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; TITLE OF INVENTION: TREATMENT OF INFLAMMATORY BOWEL DISORDERS  
; FILE REFERENCE: P1915R1 US  
; CURRENT APPLICATION NUMBER: US/10/491,997  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: PCT/US02/33070  
; PRIOR FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: US 60/340,083  
; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 162  
; SEQ ID NO 28  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-491-997-28

Query Match 100.0%; Score 1053; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.2e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDCRMARFSYVVIWMAISKVFEI GLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60  
| | | | |  
Db 1 MDCRMARFSYVVIWMAISKVFEI GLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60  
| | | | |  
Qy 61 PRSSORVPPMGIQHSKELNRTCCLLNGGTCMLGSCACPPSYGRNCEHDVRKENGCSVPH 120  
| | | | |  
Db 61 PRSSORVPPMGIQHSKELNRTCCLLNGGTCMLGSCACPPSYGRNCEHDVRKENGCSVPH 120  
| | | | |  
Qy 121 DTWLPKKCSLCKWHGQLRCFPQAF LPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180  
| | | | |  
Db 121 DTWLPKKCSLCKWHGQLRCFPQAF LPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180  
| | | | |  
Qy 181 CLSIQSY 188  
| | | | |  
Db 181 CLSIQSY 188

Search completed: February 3, 2006, 18:39:08  
Job time : 147.222 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:33:31 ; Search time 10.8622 Seconds  
(without alignments)  
202.813 Million cell updates/sec

Title: US-10-693-538-1

Perfect score: 1053

Sequence: 1 MDCRKMARFSYVIWMAIS.....SARTTTFMLVGICLSIQSY 188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:\*

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result #No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	188	6	US-10-945-853-1
2	1013	96.2	188	6	US-10-945-853-2
3	201	19.1	32	6	US-10-945-853-9
4	165	15.7	28	6	US-10-945-853-7
5	165	15.7	29	6	US-10-945-853-8
6	129	12.3	335	7	US-11-136-619-30
7	129	12.3	351	7	US-11-136-619-14
8	129	12.3	379	6	US-10-131-826A-308
9	129	12.3	379	7	US-11-136-619-2
10	129	12.3	379	7	US-11-136-619-22
11	129	12.3	379	7	US-11-136-619-23
12	121.5	11.5	1198	6	US-10-453-372-880
13	121.5	11.5	1398	6	US-10-055-877-46
14	121.5	11.5	1398	6	US-10-453-372-872
15	121.5	11.5	1403	6	US-10-055-877-52
16	121.5	11.5	1403	6	US-10-453-372-878
17	121.5	11.5	1404	6	US-10-055-877-44
18	121.5	11.5	1404	6	US-10-453-372-870
19	121.5	11.5	1450	6	US-10-055-877-48
20	121.5	11.5	1450	6	US-10-453-372-874
21	121.5	11.5	1547	6	US-10-453-372-886
22	121.5	11.5	1574	6	US-10-055-877-211
23	121.5	11.5	1577	6	US-10-055-877-54
24	121.5	11.5	1577	6	US-10-453-372-882
25	121.5	11.5	1577	6	US-10-453-372-884

Sequence 860, App  
Sequence 888, App  
Sequence 866, App  
Sequence 25, Appl  
Sequence 864, App  
Sequence 67, Appl  
Sequence 1142, Ap  
Sequence 1132, Ap  
Sequence 213, App  
Sequence 212, App  
Sequence 24, Appl  
Sequence 558, App  
Sequence 562, App  
Sequence 561, App  
Sequence 153, App  
Sequence 18, Appl  
Sequence 63, Appl  
Sequence 18, Appl  
Sequence 346, App  
Sequence 17, Appl

#### ALIGNMENTS

RESULT 1  
US-10-945-853-1  
; Sequence 1, Application US/10945853  
; Publication No. US20050255117A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc.  
; APPLICANT: Sanicola-Nadel, Michele  
; APPLICANT: Adkins, Heather  
; APPLICANT: Miklasz, Steven Donald  
; APPLICANT: Rayhorn, Paul  
; APPLICANT: Schiffer, Susan Gail  
; APPLICANT: Williams, Kevin  
; TITLE OF INVENTION: Cryptospecific Antibodies  
; FILE REFERENCE: BGN117CPCCN  
; CURRENT FILING DATE: 2004-09-20  
; PRIOR APPLICATION NUMBER: US/10/945,853  
; PRIOR APPLICATION NUMBER: PCT/US02/31462  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: PCT/US02/11950  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/286,782  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: 60/293,020  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/301,091  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/367,002  
; PRIOR FILING DATE: 2002-03-22  
; NUMBER OF SEQ ID NOS: 9  
; SEQ ID NO 1  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-945-853-1

Query Match 100.0%; Score 1053; DB 6; Length 188;  
Best Local Similarity 100.0%; Pred. No. 9.3e-100;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDCRKMARFSYVIWMAISKVFELGLVAGLGHQEFARPSRGYLAFRDDSTWPOEEPAIR 60  
Db 1 MDCRKMARFSYVIWMAISKVFELGLVAGLGHQEFARPSRGYLAFRDDSTWPOEEPAIR 60  
Qy 61 PRSSQRVPPMGIOHKSKEINRTCCCLNGGTCMLGSCACPPSPFYGRNCHEHDVKNCGSVPH 120  
Db 61 PRSSQRVPPMGIOHKSKEINRTCCCLNGGTCMLGSCACPPSPFYGRNCHEHDVKNCGSVPH 120  
Qy 121 DTWLPKCKSLCKWHGQLRCFPQAFPCDGLVMDHLVASRTPELPPSARTTTTFLMVG 180

Db 121 DTWLPKKSLCKCWHGQLRCFQAFPLPGCDGLVMDHLVASRTPELPSPSARTTTTMLVGI 180  
|  
Qy 181 CLSIQSY 188  
|  
Db 181 CLSIQSY 188  
|

## RESULT 2

US-10-945-853-2  
; Sequence 2, Application US/10945853  
; Publication No. US20050255117A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc.  
; APPLICANT: Sanicola-Nadel, Michele  
; APPLICANT: Adkins, Heather  
; APPLICANT: Miklasz, Steven Donald  
; APPLICANT: Rayhorn, Paul  
; APPLICANT: Schiffer, Susan Gail  
; APPLICANT: Williams, Kevin  
; TITLE OF INVENTION: Cripcto-Specific Antibodies  
; FILE REFERENCE: BGN117CPCCN  
; CURRENT APPLICATION NUMBER: US/10/945,853  
; CURRENT FILING DATE: 2004-09-20  
; PRIOR APPLICATION NUMBER: PCT/US02/31462  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: PCT/US02/11950  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/286,782  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: 60/293,020  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/301,091  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/367,002  
; PRIOR FILING DATE: 2002-03-22  
; NUMBER OF SEQ ID NOS: 9  
; SEQ ID NO 2  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-945-853-2

Query Match 96.2%; Score 1013; DB 6; Length 188;

Best Local Similarity 96.8%; Pred. No. 1e-95; Mismatches 6; Indels 0; Gaps 0;  
Matches 182; Conservative 0;

Qy 1 MDCRMRFSYSVIVIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDSDIWPQEEPAIR 60  
|  
Db 1 MDCRMRFSYSVIVIMAIKAFELGLVAGLGHQEFARPSRGDLAFRSDSIWPQEEPAIR 60  
|  
Qy 61 PRSSQRVPMGTOHSEKLNRTCLNGTGMGLSPCACPPSFYGRNCEHDVVRKENGCSVPH 120  
|  
Db 61 PRSSQRVPMGTOHSEKLNRTCLNGTGMGLSEFCACPPSFYGRNCEHDVVRKENGCSVPH 120  
|  
Qy 121 DTWLPKKSLCKCWHGQLRCFQAFPLPGCDGLVMDHLVASRTPELPSPSARTTTTMLVGI 180  
|  
Db 121 DTWLPKKSLCKCWHGQLRCFQAFPLPGCDGLVMDHLVASRTPELPSPSARTTTTMLAGI 180  
|  
Qy 181 CLSIQSY 188  
|  
Db 181 CLSIQSY 188  
|

## RESULT 3

US-10-945-853-9  
; Sequence 9, Application US/10945853  
; Publication No. US20050255117A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc.  
; APPLICANT: Sanicola-Nadel, Michele  
; APPLICANT: Adkins, Heather  
; APPLICANT: Miklasz, Steven Donald

; APPLICANT: Rayhorn, Paul  
; APPLICANT: Schiffer, Susan Gail  
; APPLICANT: Williams, Kevin  
; TITLE OF INVENTION: Cripcto-Specific Antibodies  
; FILE REFERENCE: BGN117CPCCN  
; CURRENT APPLICATION NUMBER: US/10/945,853  
; CURRENT FILING DATE: 2004-09-20  
; PRIOR APPLICATION NUMBER: PCT/US02/31462  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: PCT/US02/11950  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/286,782  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: 60/293,020  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/301,091  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/367,002  
; PRIOR FILING DATE: 2002-03-22  
; NUMBER OF SEQ ID NOS: 9  
; SEQ ID NO 9  
; LENGTH: 32  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: Mod\_res  
; LOCATION: 1  
; OTHER INFORMATION: N-terminal acetylation  
US-10-945-853-9

Query Match 19.1%; Score 201; DB 6; Length 32;

Best Local Similarity 96.9%; Pred. No. 2.4e-14;

Matches 31; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 119 PHDTWLPKKSLCKCWHGQLRCFQAFPLPGCD 150  
|  
Db 1 PHNTWLPKKSLCKCWHGQLRCFQAFPLPGCD 32  
|

## RESULT 4

US-10-945-853-7  
; Sequence 7, Application US/10945853  
; Publication No. US20050255117A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc.  
; APPLICANT: Sanicola-Nadel, Michele  
; APPLICANT: Adkins, Heather  
; APPLICANT: Miklasz, Steven Donald  
; APPLICANT: Rayhorn, Paul  
; APPLICANT: Schiffer, Susan Gail  
; APPLICANT: Williams, Kevin  
; TITLE OF INVENTION: Cripcto-Specific Antibodies  
; FILE REFERENCE: BGN117CPCCN  
; CURRENT APPLICATION NUMBER: US/10/945,853  
; CURRENT FILING DATE: 2004-09-20  
; PRIOR APPLICATION NUMBER: PCT/US02/31462  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: PCT/US02/11950  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/286,782  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: 60/293,020  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/301,091  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/367,002  
; PRIOR FILING DATE: 2002-03-22  
; NUMBER OF SEQ ID NOS: 9  
; SEQ ID NO 7  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:



```

; FILE REFERENCE: 023070-148711US
; CURRENT APPLICATION NUMBER: US/11/136,619
; CURRENT FILING DATE: 2005-05-23
; PRIOR APPLICATION NUMBER: US 60/664,241
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/573,197
; PRIOR FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 30
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-136-619-30

Query Match 12.3%; Score 129; DB 7; Length 335;
Best Local Similarity 30.4%; Pred. No. 5.5e-06;
Matches 31; Conservative 7; Mismatches 24; Indels 40; Gaps 5;

QY 83 CLNNGTCTMLGSCFACPPSPFYGRNCEHDVVRKENGCSVPHD---TWLPKKC----- 128
   ||| | : ||| | ||| | ||| : ||| : |||
Db 190 CMNGLCLVTPGFCICPPGFGVNCD---KANGSTTCFNNGTCTFYPGKGCICPPGLEGEQC 245
   ||| | : ||| | ||| | ||| : ||| : |||

QY 129 -----SLCKC---WHGOLRCFPQAFLPGC 149
   ||| | : ||| | |||
Db 246 EISKCPQPCRNCGKICGKSKCKSGYQGD-L-CKKPVCEPGC 286
   ||| | : ||| | |||

RESULT 7
US-11-136-619-14
; Sequence 14, Application US/11/136619
; Publication No. US2005028226A1
; GENERAL INFORMATION:
; APPLICANT: You, Liang
; APPLICANT: He, Biao
; APPLICANT: Xu, Zhidong
; APPLICANT: Jablons, David M.
; TITLE OF INVENTION: METHODS FOR TREATING AND DIAGNOSING CANCER WITH WNT INHI
; TITLE OF INVENTION: FACTOR-1 (WIP-1)
; FILE REFERENCE: 023070-148711US
; CURRENT APPLICATION NUMBER: US/11/136,619
; CURRENT FILING DATE: 2005-05-23
; PRIOR APPLICATION NUMBER: US 60/664,241
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/573,197
; PRIOR FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-136-619-14

Query Match 12.3%; Score 129; DB 7; Length 351;
Best Local Similarity 30.4%; Pred. No. 5.8e-06;
Matches 31; Conservative 7; Mismatches 24; Indels 40; Gaps 5;

QY 83 CLNNGTCTMLGSCFACPPSPFYGRNCEHDVVRKENGCSVPHD---TWLPKKC----- 128
   ||| | : ||| | ||| | ||| : ||| : |||
Db 190 CMNGLCLVTPGFCICPPGFGVNCD---KANGSTTCFNNGTCTFYPGKGCICPPGLEGEQC 245
   ||| | : ||| | ||| | ||| : ||| : |||

QY 129 -----SLCKC---WHGOLRCFPQAFLPGC 149
   ||| | : ||| | |||
Db 246 EISKCPQPCRNCGKICGKSKCKSGYQGD-L-CKKPVCEPGC 286
   ||| | : ||| | |||

RESULT 8
US-10-131-826A-308
; Sequence 308, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```



```

; APPLICANT: He, Biao
; APPLICANT: Xu, Zhidong
; APPLICANT: JiaJiong, David M.
; TITLE OF INVENTION: METHODS FOR TREATING AND DIAGNOSING CANCER WITH WNT INHIBITORY
; TITLE OF INVENTION: FACTOR-1 (WIF-1)
; FILE REFERENCE: 023070-148711US
; CURRENT APPLICATION NUMBER: US/11/136,619
; CURRENT FILING DATE: 2005-05-23
; PRIOR APPLICATION NUMBER: US 60/664,241
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/573,197
; PRIOR FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 379
; TYPE: prt
; ORGANISM: Rattus rattus
US-11-136-619-23

```

[illegible]

```

Best Local Similarity   26.1%; Pred. No. 0.00012;
Matches 49; Conservative 16; Mismatches 60; Indels 63; Gaps 10;

QY      24 ELGLVAGLGHQEPARPSRGYLAFRDSDSIWPOEIPA-IRPRS--SORVPWMGIQHSHKELNR 80
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      460 ELFPVRPLPHIAVLQDELPLQL-FQDDVDGNADEEAERLGEHTLTETKFCVLDDSFHDCSL 518
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      81 TC--CLNNGTCTML-----GSFC--ACPPSYGRNCBHDYRKEN---CGSVPHD 121
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      519 TCDRCRNGGTGLIGLDGCDCPEGWTLGLINCSCPPDTFGKNCSFSCQNGGTCDSVTGA 578
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      122 TWLP-----KKCS-----LCKWHG-----OLRCF 141
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      579 CRCPGGVSGTNCBDGCPKGYYGKHCRKCNCA NRGRCHLYGACLCDPGLYGRFCHLACP 638
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      142 POAFLPGC 149
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      639 PWAFFGPC 646
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-10-055-877-46
; Sequence 46, Application US/10055877
; Publication No. US2005028241A1
; GENERAL INFORMATION:
; APPLICANT: DeCristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tcherenev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eisen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shimkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870

```

; PRIOR FILING DATE: 2001-03-02  
 ; PRIOR APPLICATION NUMBER: 60/275,990  
 ; PRIOR FILING DATE: 2001-03-14  
 ; PRIOR APPLICATION NUMBER: 60/275,927  
 ; PRIOR FILING DATE: 2001-03-14  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 512  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 46  
 ; LENGTH: 1398  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-055-877-46

Query Match 11.5%; Score 121.5; DB 6; Length 1398;  
 Best Local Similarity 26.1%; Pred. No. 0.00014;  
 Matches 49; Conservative 16; Mismatches 60; Indels 63; Gaps 10;  
 QY 24 ELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEPA-IRPRS--SORVPPMGIQHSKELNR 80  
 DB 460 ELFPVRPLPHIAVLQDELPLQ-L-FQDDVDGADEEAEALGEHTLTKFVCLDDSFHDCSL 518  
 QY 81 TC--CLNGGTGML-----GSFC--ACPPSFYGRNCEHVDVRKEN---CGSVPHD 121  
 DB 519 TCDDCRNGGTCLLGLDGCDCPBGWTLGICNESCPDPTFGKNCSCQNGGTCDVGTGA 578  
 QY 122 TWLP-----KKCS-----LCKWHG-----QLRCF 141  
 DB 579 CRCPGVSNTCEDCGPKGYGKHKCKKNCANRGRCHRLYGACLCDFGLYGRFCHLACP 638  
 QY 142 PQAFPLPGC 149  
 DB 639 PWAFGPGC 646

RESULT 14  
 US-10-453-372-872  
 ; Sequence 872, Application US/10453372  
 ; Publication No. US2006000323A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alsobrook, Therapeutics, Inc.  
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
 ; FILE REFERENCE: 21402-589 A  
 ; CURRENT APPLICATION NUMBER: US/10/453,372  
 ; CURRENT FILING DATE: 2003-06-03  
 ; PRIOR APPLICATION NUMBER: 09/789390  
 ; PRIOR FILING DATE: 2001-02-23  
 ; PRIOR APPLICATION NUMBER: 60/185967  
 ; PRIOR FILING DATE: 2000-03-01  
 ; PRIOR APPLICATION NUMBER: 09/823187  
 ; PRIOR FILING DATE: 2001-03-29  
 ; PRIOR APPLICATION NUMBER: 60/195792  
 ; PRIOR FILING DATE: 2000-03-10  
 ; PRIOR APPLICATION NUMBER: 09/839446  
 ; PRIOR FILING DATE: 2001-03-19  
 ; PRIOR APPLICATION NUMBER: 60/199476  
 ; PRIOR FILING DATE: 2000-03-25  
 ; PRIOR APPLICATION NUMBER: 09/863776  
 ; PRIOR FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: 60/208263  
 ; PRIOR FILING DATE: 2000-05-31  
 ; PRIOR APPLICATION NUMBER: 09/939398  
 ; PRIOR FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: 60/227800  
 ; PRIOR FILING DATE: 2000-08-25  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 1609  
 ; SOFTWARE: CuraSeqList version 0.1  
 ; SEQ ID NO 872  
 ; LENGTH: 1398  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-453-372-872

Query Match 11.5%; Score 121.5; DB 6; Length 1398;  
 Best Local Similarity 26.1%; Pred. No. 0.00014;  
 Matches 49; Conservative 16; Mismatches 60; Indels 63; Gaps 10;  
 QY 24 ELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEPA-IRPRS--SORVPPMGIQHSKELNR 80  
 DB 460 ELFPVRPLPHIAVLQDELPLQ-L-FQDDVDGADEEAEALGEHTLTKFVCLDDSFHDCSL 518  
 QY 81 TC--CLNGGTGML-----GSFC--ACPPSFYGRNCEHVDVRKEN---CGSVPHD 121  
 DB 519 TCDDCRNGGTCLLGLDGCDCPBGWTLGICNESCPDPTFGKNCSCQNGGTCDVGTGA 578  
 QY 122 TWLP-----KKCS-----LCKWHG-----QLRCF 141  
 DB 579 CRCPGVSNTCEDCGPKGYGKHKCKKNCANRGRCHRLYGACLCDFGLYGRFCHLACP 638  
 QY 142 PQAFPLPGC 149  
 DB 639 PWAFGPGC 646

RESULT 15  
 US-10-055-877-52  
 ; Sequence 52, Application US/10055877  
 ; Publication No. US2005028241A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DeCristofaro, Marc  
 ; APPLICANT: Padigaru, Murallidhara  
 ; APPLICANT: Miller, Charles  
 ; APPLICANT: Tchernev, Velizar  
 ; APPLICANT: Zhong, Mei  
 ; APPLICANT: Anderson, David  
 ; APPLICANT: Ballinger, Robert  
 ; APPLICANT: Gerlach, Valerie  
 ; APPLICANT: Spytek, Kimberly  
 ; APPLICANT: Ratelli, Luca  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Guo, Xiaojia  
 ; APPLICANT: Zerhusen, Bryan  
 ; APPLICANT: Andrew, David  
 ; APPLICANT: Mezes, Peter  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Burgess, Catherine  
 ; APPLICANT: Eisen, Andrew  
 ; APPLICANT: Wolenc, Adam  
 ; APPLICANT: Baumgartner, Jason  
 ; APPLICANT: Shimkets, Richard  
 ; APPLICANT: Gusev, Vladimir  
 ; APPLICANT: Vernet, Corine  
 ; APPLICANT: Taupier Jr., Raymond  
 ; APPLICANT: Pena, Carol  
 ; APPLICANT: Shenoy, Suresh  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Casman, Stacie  
 ; APPLICANT: Boldog, Ference  
 ; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby  
 ; FILE REFERENCE: 21402-251  
 ; CURRENT APPLICATION NUMBER: US/10/055,877  
 ; CURRENT FILING DATE: 2002-01-22  
 ; PRIOR APPLICATION NUMBER: 60/262,892  
 ; PRIOR FILING DATE: 2001-01-19  
 ; PRIOR APPLICATION NUMBER: 60/263,598  
 ; PRIOR FILING DATE: 2001-01-23  
 ; PRIOR APPLICATION NUMBER: 60/263,799  
 ; PRIOR FILING DATE: 2001-01-24  
 ; PRIOR APPLICATION NUMBER: 60/264,117  
 ; PRIOR FILING DATE: 2001-01-25  
 ; PRIOR APPLICATION NUMBER: 60/264,139  
 ; PRIOR FILING DATE: 2001-01-25  
 ; PRIOR APPLICATION NUMBER: 60/264,478  
 ; PRIOR FILING DATE: 2001-01-26  
 ; PRIOR APPLICATION NUMBER: 60/263,351

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
; LENGTH: 1403
; TYPE: EXT
; ORGANISM: Homo sapiens
US-10-055-877-52

Query Match      11.5%; Score 121.5; DB 6; Length 1403;
Best Local Similarity 26.1%; Pred. No. 0.00014;
Matches 49; Conservative 16; Mismatches 60; Indels 63; Gaps 10;

Qy      24  ELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPA-IRPRS--SQRVPPMGIQHSKEINR 80
Db      466  ELPPVRPLPHIAVLQDELPLQ-L-FQDDDVGADEEAELRGEHTLTEKFCVCLDDSFHDCSL 524

Qy      81  TC--CLNGGTGML-----GSFC--ACPPSFYGRNCEHVRKEN---CGSVPHD 121
Db      525  TCDDCRNGGTCLLGLDGDCEPGWTGLICNESCPDPTFGKNCSPSCSCQNGGTCDSVTGA 584

Qy      122  TWLP-----KKCS-----LCKCWHG-----QLRCF 141
Db      585  CRCPPGVSGTNCEDGCPKGYGKHCRKCKNCANRGRCHRLYGACLCDPGLYGRFCHLACP 644

Qy      142  PQAFLEPGC 149
Db      645  PWAFPGGC 652
```

Search completed: February 3, 2006, 18:39:40  
Job time : 11.8622 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:23:01 ; Search time 30.9156 Seconds  
(without alignments)  
585.102 Million cell updates/sec

Title: US-10-693-538-1  
Perfect score: 1053  
Sequence: 1 MDCRKMARSYSYVIMTMAIS.....SARTTFMLVGICLSIQSY 188  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	188	2 A30362	teratocarcinoma-de
2	1013	96.2	188	2 A39787	teratocarcinoma-de
3	647	61.4	171	2 I49612	teratocarcinoma-de
4	129	12.3	379	2 A59180	Wnt inhibitory fac
5	125.5	11.9	4351	2 T00252	MEGF1 protein - ra
6	121.5	11.5	1574	2 T13954	MEGF6 protein - ra
7	121	11.5	378	2 B59180	Wnt inhibitory fac
8	119	11.3	2555	2 A40043	notch protein homo
9	118.5	11.3	1722	2 B89753	protein Fl1C7.4 (i
10	117.5	11.2	477	1 A34369	t-plasminogen acti
11	117.5	11.2	477	2 J05098	t-plasminogen acti
12	115.5	11.0	2437	2 S42612	transmembrane prot
13	115.5	11.0	2524	2 A35844	Xotch protein - Af
14	115	10.9	2471	2 A49128	cell-face determin
15	114	10.8	387	2 B49175	Notch A protein -
16	114	10.8	861	2 A48825	Notch homolog Motc
17	114	10.8	2531	2 S18188	notch protein homo
18	114	10.8	2531	2 A46019	notch-1 protein -
19	113	10.7	427	2 J04915	ags protein precu
20	111	10.5	477	2 J05097	t-plasminogen acti
21	110.5	10.5	559	1 A29941	t-plasminogen acti
22	110	10.4	835	2 J00076	nel protein - chic
23	108.5	10.3	1203	2 A49175	Notch B protein -
24	108.5	10.3	1620	2 T27283	hypothetical prote
25	108	10.3	293	2 T09065	hypothetical prote
26	107	10.2	686	2 J07569	Delta-4 protein -
27	106.5	10.1	291	2 I38098	t-plasminogen acti
28	106.5	10.1	562	1 UKHUT	t-plasminogen acti
29	106	10.1	559	1 A35029	t-plasminogen acti

30 103.5 9.8 1125 1 S57846 protein-tyrosine k  
31 102.5 9.7 2703 1 A24420 notch protein - fr  
32 101.5 9.6 640 2 T08179 LRG5 protein - chl  
33 101.5 9.6 2139 2 A35672 crumbe protein - f  
34 101.5 9.6 3871 2 T22812 hypothetical prote  
35 101 9.6 685 2 J07570 Delta-4 protein -  
36 101 9.6 1429 2 S06434 homeotic protein l  
37 100 9.5 293 2 B26637 neurogenic repetit  
38 100 9.5 2321 2 S78549 notch3 protein - h  
39 99.5 9.4 1064 2 A40136 fibropellin Ia - s  
40 99.5 9.4 2352 2 T30201 Notch homolog prot  
41 99 9.4 1124 1 I58388 Notch tyrosine k  
42 99 9.4 1964 2 T09059 notch4 - mouse  
43 98 9.3 593 2 S45281 coagulation factor  
44 98 9.3 1136 1 S57845 protein-tyrosine k  
45 98 9.3 1138 1 S24066 protein-tyrosine k

ALIGNMENTS

RESULT 1

A30362  
teratocarcinoma-derived growth factor 1 - human  
N;Alternate names: CRIPTO protein  
C;Species: Homo sapiens (man)  
C;Date: 18-Oct-1989 #sequence\_revision 18-Oct-1989 #text\_change 09-Jul-2004  
C;Accession: B39787; A30362  
R;Dono, R.; Montuori, N.; Rocchi, M.; De Ponti-Zilli, L.; Ciccodicola, A.; Persico, M.G.  
Am. J. Hum. Genet. 49, 555-565, 1991  
A;Title: Isolation and characterization of the CRIPTO autosomal gene and its X-linked re  
A;Reference number: A39787; MUID:91353571; PMID:1882841  
A;Accession: B39787  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-188 <DON>  
A;Cross-references: UNIPARC:P13385; UNIPARC:UPI000004966D; GB:M96955; GB:M37099; NID:G33-  
EMBO J. 8, 1987-1991, 1989  
R;Ciccodicola, A.; Dono, R.; Obici, S.; Simeone, A.; Zollo, M.; Persico, M.G.  
A;Title: Molecular characterization of a gene of the 'EGF family' expressed in undifferent  
A;Reference number: A30362; MUID:90005403; PMID:2792079  
A;Accession: A30362  
A;Molecule type: mRNA  
A;Residues: 1-188 <CIC>  
A;Cross-references: UNIPARC:UPI000004966D; GB:X14253; NID:G30220; PIDN:CAA32467.1; PID:G  
C;Superfamily: teratocarcinoma-derived growth factor 1; EGF homology  
C;Keywords: Growth factor  
F;78-106/Domain: EGF homology <EGF>

Query Match 100.0%; Score 1053; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1e-89;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDCRKMARSYSYVIMTMAISKVPELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
Db 1 MDCRKMARSYSYVIMTMAISKVPELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
Qy 61 PRSSQRVPMGIQHSKELNRTCCINGTCMLGSCFACPPSFYGRNCEHDVKNCGSVPH 120  
Db 61 PRSSQRVPMGIQHSKELNRTCCINGTCMLGSCFACPPSFYGRNCEHDVKNCGSVPH 120  
Qy 121 DTWLPPKCSLCKWHQQLRCFFQAFPLFGCDGLVMDHLVASRTPPELPPSARTTTFMVLGI 180  
Db 121 DTWLPPKCSLCKWHQQLRCFFQAFPLFGCDGLVMDHLVASRTPPELPPSARTTTFMVLGI 180  
Qy 181 CLSIQSY 188  
Db 181 CLSIQSY 188

RESULT 2

A39787  
teratocarcinoma-derived growth factor - human

C:Species: Homo sapiens (man)  
 C:Date: 20-Mar-1992 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
 C:Accession: A39787  
 R:Dono, R.; Montuori, N.; Rocchi, M.; De Ponti-Zilli, L.; Ciccodicola, A.; Persico, M.G.; Am. J. Hum. Genet. 49, 555-565, 1991  
 A:Title: Isolation and characterization of the CRPTO autosomal gene and its X-linked re  
 A:Reference number: A39787; MUID:91353571; PMID:1882841  
 A:Accession: A39787  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-188 <DON>  
 A:Cross-references: UNIPROT:Q8TCC1; UNIPARC:UPI0000145185  
 A:Note: the authors translated the codon GAC for residue 43 as Glu  
 C:Superfamily: teratocarcinoma-derived growth factor 1; EGF homology

Query Match 96.2%; Score 1013; DB 2; Length 188;  
 Best Local Similarity 96.8%; Pred. No. 5e-86;  
 Matches 182; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MDCRWARFSYSVIMAIKVFELGLVAGLGHQEFARPSRGYLAFRRDSDIWPQEEPAIR 60  
 DB 1 MDCRWARFSYSVIMAIKAFELGLVAGLGHQEFARPSRGDLAFRRDSDIWPQEEPAIR 60  
 QY 61 PRSSQRVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVRKENGCSVPH 120  
 DB 61 PRSSQRVPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVRKENGCSVPH 120  
 QY 121 DTWLPKCSLCKWHGQLRCPQAFPGCDGLVMDHVLVASRTPELPPSARTTTFMLVGI 180  
 DB 121 DTWLPKCSLCKWHGQLRCPQAFPGCDGLVMDHVLVASRTPELPPSARTTTFMLAGI 180  
 QY 181 CLSIQSY 188  
 DB 181 CLSIQSY 188

RESULT 3  
 I49612  
 teratocarcinoma-derived growth factor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
 C:Accession: I49612  
 R:Dono, R.; Scalera, L.; Pacifico, F.; Simeone, A.; Persico, M.G.; Acampora, D.  
 Development 118, 1157-1168, 1993  
 A:Title: The murine cripto gene: expression during mesoderm induction and early heart m  
 A:Reference number: I49612; MUID:94094736; PMID:7916676  
 A:Accession: I49612  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-171 <RES>  
 A:Cross-references: UNIPROT:P51865; UNIPARC:UPI0000029326; GB:M87321; NID:9402714; PIDN:  
 C:Genetics:  
 A:Gene: cripto  
 C:Superfamily: teratocarcinoma-derived growth factor 1; EGF homology

Query Match 61.4%; Score 647; DB 2; Length 171;  
 Best Local Similarity 67.8%; Pred. No. 2.2e-52;  
 Matches 120; Conservative 14; Mismatches 31; Indels 12; Gaps 2;

QY 6 MARFSYSVIMAIKVFELGLVAGLGHQEFARPSRGYLAFRRDSDIWPQEEPAIRPRGSQ 65  
 DB 1 MCFSSSVLLVAISAFEGFVAGRD-----LAIRDNIWQKEPAVDRDFQ 49  
 QY 66 RYPPMGIQHSKELNRTCCCLNGGTCMLGSFCACPPSFYGRNCEHDVRKENGCSVPHDTWLP 125  
 DB 50 FVPSVGIQNSKLNKTCCLNGGTCILGSFCACPPSFYGRNCEHDVRKENGCSILHGTWLP 109  
 QY 126 KCSLCKWHGQLRCPQAFPGCDGLVMDHVLVASRTPELPPSARTTTFMLVGLCL 182  
 DB 110 KCSLCKWHGQLHCLPQTFPLGCDGHVMDQLKASRTPCQTPSV-TTTFMLAGACL 165

RESULT 4

A59180  
 Wnt inhibitory factor-1 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 09-Jul-2004  
 C:Accession: A59180  
 R:Hsieh, J.C.; Kodjabachian, L.; Rebbert, M.L.; Rattner, A.; Smallwood, P.M.; Samos, C.H.  
 Nature 398, 431-436, 1999  
 A:Title: A new secreted protein that binds to Wnt proteins and inhibits their activities  
 A:Reference number: A59180; MUID:99215557; PMID:10201374  
 A:Accession: A59180  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-379 <HSI>  
 A:Cross-references: UNIPROT:Q9Y5W5; UNIPARC:UPI0000051058; GB:AF122922; NID:94585369; PI

Query Match 12.3%; Score 129; DB 2; Length 379;  
 Best Local Similarity 30.4%; Pred. No. 0.00021;  
 Matches 31; Conservative 7; Mismatches 24; Indels 40; Gaps 5;

QY 83 CLNGGTCMLGSFCACPPSFYGRNCEHDVRKENGCSVPHD---TWLPKKC----- 128  
 DB 218 CMNGGLCTVPGFCICPPGFYGVNCD----XANCSTTCFNGGTCFYFGKICICPPGLEGEQC 273  
 QY 129 -----SLCKC---WHGQLRCPQAFPLPGC 149  
 DB 274 EISKCPQCRNGGKICGKCKCKSGYQGD-L-CKKPCEPGC 314

RESULT 5  
 T00252  
 MEGF1 protein - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: T00252  
 R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
 Genomics 51, 27-34, 1998  
 A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
 A:Reference number: 214126; MUID:98360089; PMID:9693030  
 A:Accession: T00252  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-4351 <NAK>  
 A:Cross-references: UNIPROT:O88277; UNIPARC:UPI0000055B23; EMBL:AB011527; NID:93449285;  
 A:Experimental source: brain  
 C:Genetics:  
 A:Gene: MEGF1  
 C:Superfamily: rat MEGF1 protein; cadherin repeat homology; EGF homology; laminin G repe  
 F:3798-3949/Domain: laminin G repeat homology <LGR>  
 F:3953-3985/Domain: EGF homology <EGF>  
 F:3992-4023/Domain: EGF homology <EGF1>

Query Match 11.9%; Score 125.5; DB 2; Length 4351;  
 Best Local Similarity 30.9%; Pred. No. 0.0041;  
 Matches 43; Conservative 14; Mismatches 29; Indels 53; Gaps 11;

QY 63 SSQRVPPMGIQHSKE-----LNRTC-----CLNGGTC--MLGS--FCACPPSPF 101  
 DB 3921 NGERLELLIGREKKEGRLETWALSOCWPGFACQSQCPLNGGSCSPALGSLVLCCKCPDPF 3980  
 QY 102 YGRNCEHDVRKENGCSVPHDTWLPKKC-----SLCKCWHGQL--RCFPQAFPLP 147  
 DB 3981 SGRNCE--LGRNCTSAP-----CQEGGTCVSSPSGTCNCNCPHYTGDRCEMA--R 4028  
 QY 148 GCDGLVMDHVLVASRTPEL 166  
 DB 4029 GCSG----GHCLLI--TPEI 4041

RESULT 6  
 T13954  
 MEGF6 protein - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004



C;Accession: T13954  
R;Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
Genomics 51, 27-34, 1998  
A;Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
A;Reference number: 214126; MUID:98360089; PMID:9693030  
A;Accession: T13954  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1574 <NAK>  
A;Cross-references: UNIPROT:O88281; UNIPARC:UPI0000043BEE; EMBL:AB011532; NID:g3449293;  
A;Experimental source: strain Sprague-Dawley; brain  
C;Genetics:  
A;Gene: MEGF6

Query Match 11.5%; Score 121.5; DB 2; Length 1574;  
Best Local Similarity 30.3%; Pred. No. 0.0038;  
Matches 37; Conservative 11; Mismatches 31; Indels 43; Gaps 10;

Qy 57 PAIR--PRSSQVPM--GIQHSKELNRTC--CLNGGTC-----MLGSFC--AC 97  
Db 982 PAGRWGPRCAQSCPLPTFGLNCS----QICTCFNGASCDSVTGQCHCAPGMGPTCLQAC 1037

Qy 98 PPSFYGRNCEDHVRKNCVPHDTWLPKKC---SLCKC---WHG---QLRCFPQAFPLP 147  
Db 1038 PPGLYGKNCQHSCLCRNGG-----RCDPILGQCTCPGWTGLACENECLPGHYAA 1087

Qy 148 GC 149  
Db 1088 GC 1089

RESULT 7  
B59180  
Wnt inhibitory factor-1 - zebra fish  
C;Species: Brachydanio rerio (zebra fish)  
C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 09-Jul-2004  
C;Accession: B59180  
R;Haieh, J.C.; Kodjabachian, L.; Rebbert, M.L.; Rattner, A.; Smallwood, P.M.; Samos, C.H.  
Nature 398, 431-436, 1999  
A;Title: A new secreted protein that binds to Wnt proteins and inhibits their activities  
A;Reference number: A59180; MUID:99215557; PMID:10201374  
A;Accession: B59180  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-378 <HGI>  
A;Cross-references: UNIPROT:Q9W6F9; UNIPARC:UPI0000138F08; GB:AF122925; NID:g4585375; PI  
C;Genetics:  
A;Gene: WIF-1

Query Match 11.5%; Score 121; DB 2; Length 378;  
Best Local Similarity 28.0%; Pred. No. 0.0011;  
Matches 33; Conservative 8; Mismatches 29; Indels 48; Gaps 6;

Qy 71 GIQHSKELNRTCCLNGGTCMLGSCACPPSFYGRNCEDHVRKNC-----GSVPHDTW 123  
Db 204 GVHCEKALCSPRLNGGLCMSPGVICPPGFGSSCE-----RANCSITCLNGGTCFPH--- 256

Qy 124 LPKKC-----SLCKC---WHGQLRCFPQAFPLGC 149  
Db 257 -PGKICAVSFEVGRCELSKRCQPCRNKGKCTGRNCKCKSGYHGDL-CKSAVCEPSC 312

RESULT 8  
A40043  
notch protein homolog TAN-1 precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 21-Apr-1992 #sequence\_revision 21-Apr-1992 #text\_change 05-Oct-2004  
C;Accession: A40043  
R;Ellisen, L.W.; Bird, J.; West, D.C.; Soreng, A.L.; Reynolds, T.C.; Smith, S.D.; Sklar, C.  
Cell 66, 649-661, 1991  
A;Title: TAN-1, the human homolog of the Drosophila Notch gene, is broken by chromosomal  
A;Reference number: A40043; MUID:91347367; PMID:1831692  
A;Accession: A40043

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
A;Molecule type: mRNA  
A;Residues: 1-2555 <ELL>  
A;Cross-references: UNIPARC:UPI0000177455; GB:M73980  
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology  
F;261-252/Domain: EGF homology <EGX1>  
F;494-525/Domain: EGF homology <EGF1>  
F;987-1018/Domain: EGF homology <EGX2>  
F;1149-1180/Domain: EGF homology <EGF>  
F;1187-1218/Domain: EGF homology <EGF3>  
F;1233-1264/Domain: EGF homology <EGX3>  
F;1927-1959/Domain: ankyrin repeat homology <AN1>  
F;1960-1992/Domain: ankyrin repeat homology <AN2>  
F;1994-2026/Domain: ankyrin repeat homology <AN3>  
F;2027-2059/Domain: ankyrin repeat homology <AN4>  
F;2060-2092/Domain: ankyrin repeat homology <AN5>

Query Match 11.3%; Score 119; DB 2; Length 2555;  
Best Local Similarity 36.4%; Pred. No. 0.01;  
Matches 32; Conservative 10; Mismatches 28; Indels 18; Gaps 7;

Qy 74 HSKELNRTCCLNGGTCM--LGSP--CACPPSFYGRNCEDHVRKNCVPHDTWLPKCKSL 130  
Db 983 NTPDCTSSCFENGCTVDGINSFTCLCPPGFTGSGYQHVNE--CDSRP-----CLL 1032

Qy 131 ---CKWHGQLRC-FPOAFL-PCDGLV 153  
Db 1033 GGTCDGGRGLHRCCTCPQGTGPNCONLV 1060

RESULT 9  
E89753  
protein Flt1C7.4 [imported] - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C;Accession: E89753  
R;anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A;Reference number: A75000; MUID:99069613; PMID:9851916  
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.ele  
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A;Accession: E89753  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1722 <STO>  
A;Cross-references: UNIPROT:Q19350; UNIPARC:UPI0000074975; GB:chr\_X; PIDN:AAC69012.1; PI  
C;Genetics:  
A;Gene: Flt1C7.4  
A;Map position: X

Query Match 11.3%; Score 118.5; DB 2; Length 1722;  
Best Local Similarity 25.9%; Pred. No. 0.0078;  
Matches 36; Conservative 12; Mismatches 42; Indels 49; Gaps 6;

Qy 79 NRTC-----CLNGGTCMLGS-----FCACPPSFYGRNCEDHVRKNC-----GSVPHDTW 123  
Db 19 NRACSRNTCLNGGTCVTNDTRMFQCECPKGFSGLLCQ-----DNCSLHCLHGNVCVGTG 73

Qy 124 LPKKCSLCKWHGQLRCFPQAFPLPGCDGLVNDHLVASRTPEL----- 166  
Db 74 GEETCQCEGMGSL-----CDNLVTDDDTAQKCSFQCGDDBRCTKGADGSYICQ 123

Qy 167 --PPSARTTTTFLVGLICLS 183  
Db 124 SNEPSCAHTCQNGTCTVA 142

RESULT 10  
A34369  
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)  
C;Species: Megaderma lyra  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

Db 17 GAVFSLPQETRYQLARGSRAYGVACRDEKTMQYQQQBSWLRPEVRSKRK-----EHCR 71  
 QY 76 -----KELNRTCCLLNGGTC-MLGSF-----CACPPSFYGRNCEHVDVRK---EN 114  
 Db 72 CDREGLAQCHTVPVKSCSELRCFNGGTCWQAASFDFVCCQPKGYTGKQCEVDTHATCYKD 131  
 QY 115 CGSVPHDTWLPKKC-SLCKWHGQL 138  
 Db 132 QGVYRGITWSTSESQAQCINWNSNL 156

RESULT 12  
 S42612  
 transmembrane protein precursor - zebra fish  
 C:Species: Brachydanio rerio (zebra fish)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
 C:Accession: S42612  
 R;Bierkamp, C.; Campos-Ortega, J.A.  
 Mech. Dev. 43, 87-100, 1993  
 A;Title: A zebrafish homologue of the Drosophila neurogenic gene Notch and its  
 A;Reference number: S42612; MUID:94128602; PMID:8297791  
 A;Accession: S42612  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-2437 <BI>  
 A;Cross-references: UNIPROT:P46530; UNIPARC:UPI0000130565; EMBL:X69088; NID:943  
 C;Superfamily: notch protein; ankyrin repeat homology; EGF homology  
 P;755-786/Domain: EGF homology <EGF1>  
 F;1023-1034/Domain: EGF homology <EGF>  
 F;1185-1216/Domain: EGF homology <EGF2>  
 F;1915-1947/Domain: ankyrin repeat homology <AN1>  
 F;1948-1980/Domain: ankyrin repeat homology <AN2>  
 F;1982-2014/Domain: ankyrin repeat homology <AN3>  
 F;2015-2047/Domain: ankyrin repeat homology <AN4>  
 F;2048-2080/Domain: ankyrin repeat homology <AN5>

Query Match 11.0%; Score 115.5; DB 2; Length 2437;  
 Best Local Similarity 34.8%; Pred. No. 0.02;  
 Matches 31; Conservative 12; Mismatches 27; Indels 19; Gaps 7;

QY 74 HSKELNRTCLNGGTCM-LGSF-CACPPSFYGRNCEHVDVRKNCGSVP-----HDTW 123  
 Db 981 NTPDCTESSCFNGGTCVDGISSFCVCLPFTGNYCQHDVNE--CDSRPCQNGSGSCDGY 1038

QY 124 LPKKSILCKQWGG--LRCPQAFPLPGCD 150  
 Db 1039 GTYKCT---CPHYTGLNC--QSLVRWCD 1062

RESULT 13  
 A35844  
 Notch protein - African clawed frog  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 12-Oct-1990 #sequence\_revision 12-Oct-1990 #text\_change 05-Oct-2004  
 C:Accession: A35844  
 R;Coffman, C.; Harris, W.; Kintner, C.  
 Science 249, 1438-1441, 1990  
 A;Title: Notch, the Xenopus homolog of Drosophila notch.  
 A;Reference number: A35844; MUID:90385285; PMID:2402639  
 A;Accession: A35844  
 A;Status: preliminary; nucleic acid sequence not shown; not compared with concen  
 A;Molecule type: mRNA  
 A;Residues: 1-2524 <COF>  
 A;Cross-references: UNIPARC:UPI000004F253  
 C;Superfamily: notch protein; ankyrin repeat homology; EGF homology  
 C;Keywords: transmembrane protein  
 F;146-177/Domain: EGF homology <EGX1>  
 F;184-215/Domain: EGF homology <EGF1>  
 F;222-254/Domain: EGF homology <EGF>  
 F;456-487/Domain: EGF homology <EGX2>  
 F;757-788/Domain: EGF homology <EGF3>  
 F;1025-1056/Domain: EGF homology <EGX3>

F;1924-1956/Domain: ankyrin repeat homology <AN1>  
F;1957-1989/Domain: ankyrin repeat homology <AN2>  
F;1991-2023/Domain: ankyrin repeat homology <AN3>  
F;2024-2056/Domain: ankyrin repeat homology <AN4>  
F;2057-2089/Domain: ankyrin repeat homology <AN5>

Query Match 11.0%; Score 115.5; DB 2; Length 2524;  
Best Local Similarity 32.2%; Pred. No. 0.021;  
Matches 28; Conservative 15; Mismatches 29; Indels 15; Gaps 6;

Qy 74 HSKELNRTCCLLNGGTCM--LGSP-CACPPSPVGRNCEHDVRKENGSGVP-----HDTW 123  
Db 983 NTPDCTESSFCNGTCDINGINTFCQPPGFTGSCQHDINE--CDSPCLNGGTCDQSY 1040  
Qy 124 LPKCSLKCKWHGQLRCPQAFPLPGCD 150  
Db 1041 GTYRKT-CPQGYTGLNC--QNLVRWCD 1064

RESULT 14  
A49128  
cell-fate determining gene Notch2 protein - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 05-Oct-2004  
C;Accession: A49128  
R;Weinmaster, G.; Roberts, V.J.; Lemke, G.  
Development 116, 931-941, 1992  
A;Title: Notch2: a second mammalian Notch gene.  
A;Reference number: A49128; MUID:93202015; PMID:1295745  
A;Accession: A49128  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-2471 <WEI>  
A;Cross-references: UNIPROT:Q90W30; UNIPARC:UPI000004F255  
A;Experimental source: Schwann cell  
A;Note: sequence extracted from NCBI backbone (NCBIP:127811)  
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology  
F;264-295/Domain: EGF homology <EGX1>  
F;799-830/Domain: EGF homology <EGF1>  
F;877-908/Domain: EGF homology <EGX2>  
F;1029-1060/Domain: EGF homology <EGF>  
F;1067-1098/Domain: EGF homology <EGX3>  
F;1153-1184/Domain: EGF homology <EGF3>  
F;1191-1222/Domain: EGF homology <EGX4>  
F;1876-1908/Domain: EGF homology <EGX4>  
F;1909-1941/Domain: ankyrin repeat homology <AN1>  
F;1943-1975/Domain: ankyrin repeat homology <AN2>  
F;1976-2008/Domain: ankyrin repeat homology <AN3>  
F;2009-2041/Domain: ankyrin repeat homology <AN4>  
F;2009-2041/Domain: ankyrin repeat homology <AN5>

Query Match 10.9%; Score 115; DB 2; Length 2471;  
Best Local Similarity 30.2%; Pred. No. 0.023;  
Matches 38; Conservative 6; Mismatches 46; Indels 36; Gaps 8;

Qy 83 CLNGGTCLM-----GSFCACPPSPFYGRNCEHDV-----RKNC---GSVPDHTWLPKK 127  
Db 1313 CLNGGTCAVASNPVDGFTICRCPGFGSGARCGSCGVKCRGECQVHTASGP-FCFCPNH 1371  
Qy 128 -----CSLCKCKWHGQLRCPQAFPLP-----GCDGLVDEHLVASRTPELPSPARTTTFML 177  
Db 1372 KCESGCASNPQRGG-TCYPQRPYPYSCRSPFPWGSCHESYTAPTSTPPA----- 1423  
Qy 178 VGICLS 183  
Db 1424 --TCLS 1427

RESULT 15  
B49175  
Motch A protein - mouse (fragment)  
N;Alternate names: Notch homolog  
C;Species: Mus musculus (house mouse)  
C;Date: 21-Jan-1994 #sequence\_revision 05-Jan-1996 #text\_change 05-Oct-2004

C;Accession: B49175; PH1569; S32109  
R;Lardelli, M.; Lendahl, U.  
Exp. Cell Res. 204, 364-372, 1993  
A;Title: Motch A and Motch B--two mouse Notch homologues coexpressed in a wide variety o  
A;Reference number: A49175; MUID:93178563; PMID:8440332  
A;Accession: B49175  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-387 <LAR>  
A;Cross-references: UNIPROT:Q01705; UNIPARC:UPI000016CF1E; EMBL:X68278; NID:g287987; PID  
A;Experimental source: embryo  
A;Note: sequence extracted from NCBI backbone (NCBIP:126159)  
C;Comment: This protein has many EGF repeats and lin-12/Notch repeats.  
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology  
F;27-58/Domain: EGF homology <EGF>  
F;73-104/Domain: EGF homology <EGF2>  
F;151-185/Domain: EGF homology <EGF1>

Query Match 10.8%; Score 114; DB 2; Length 387;  
Best Local Similarity 29.6%; Pred. No. 0.0052;  
Matches 34; Conservative 11; Mismatches 34; Indels 36; Gaps 7;

Qy 49 DSTWPQEEPAIR-PRSSQRPVPMGIGHKELNRTCCLLNGGTCM--LGSP-CACPPSPFYGR 104  
Db 63 DDCHPPLDPASRSPK-----CFNNGTCDVQGVGYTCTCPPGFVGE 102  
Qy 105 NCEHDVRKENGSGVPDHTWLPKKC-----SLCKCKWHGQLRCPQAFPLPGCDG-151  
Db 103 RCEGDVNE--CLSNPCDPRGTQNCVQRVNDPHCECRAGHTGRRRC--ESVINGCRG 153

Search completed: February 3, 2006, 18:31:16  
Job time : 31.9156 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:15:30 ; Search time 181.316 Seconds  
(without alignments)  
731.539 Million cell updates/sec

Title: US-10-693-538-1

Perfect score: 1053

Sequence: 1 MDCRMARFSYVIMWMAIS.....SARTTTPMLVIGICLSIQSY 188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	188	1 TDGF1_HUMAN	P13385 homo sapien
2	1049	99.6	188	2 Q8TCC1_HUMAN	Q8tcc1 homo sapien
3	1013	96.2	188	1 TDGF2_HUMAN	P51864 homo sapien
4	790	75.0	181	2 Q58D57_BOVIN	Q58d57 bos taurus
5	647	61.4	171	1 TDGF1_MOUSE	P51865 mus musculus
6	640	60.8	171	2 Q7IQ06_MOUSE	Q7iq06 mus musculus
7	363	34.5	66	2 Q9IAT2_BRARE	Q9iat2 brachydanio
8	274.5	26.1	183	2 Q57517_BRARE	O57517 brachydanio
9	274.5	26.1	183	2 Q50415_BRARE	Q50415 brachydanio
10	274.5	26.1	190	2 Q57516_BRARE	O57516 brachydanio
11	274.5	26.1	190	2 Q91649_XENLA	Q91649 xenopus lae
12	268	25.5	190	2 Q80042_XENLA	Q80042 xenopus lae
13	256.5	24.4	193	2 Q91803_CHICK	Q91803 gallus gall
14	233	22.1	223	2 Q53T05_HUMAN	Q53t05 homo sapien
15	233	22.1	223	2 Q9GZR3_HUMAN	Q9gzr3 homo sapien
16	225.5	21.4	202	2 P97766_MOUSE	P97766 mus musculus
17	197	18.7	154	2 Q9JIB7_MOUSE	Q9jib7 mus musculus
18	129	12.3	220	2 Q99KR2_MOUSE	Q99kr2 mus musculus
19	129	12.3	379	1 W1F1_HUMAN	Q9y5w5 homo sapien
20	129	12.3	379	1 W1F1_MOUSE	Q9wual mus musculus
21	129	12.3	379	1 W1F1_RAT	Q6in38 rattus norv
22	128.5	12.2	1654	2 Q69ZV7_MOUSE	Q69zv7 mus musculus
23	128.5	12.2	4351	2 Q5F226_MOUSE	Q5f226 mus musculus
24	127	12.1	2067	2 Q59ED8_HUMAN	Q59ed8 homo sapien
25	127	12.1	2555	2 Q5SXN3_HUMAN	Q5sxm3 homo sapien
26	125.5	11.9	4351	1 FAT2_RAT	Q88277 rattus norv
27	123	11.7	279	1 E6GLF7_RAT	Q6az60 rattus norv
28	123	11.7	1721	2 Q614N6_CAEBR	Q614n6 caenorhabdi
29	121.5	11.5	1574	1 E6FLF7_RAT	Q88281 rattus norv
30	121	11.5	378	1 W1F1_BRARE	Q9w6f9 brachydanio
31	121	11.5	378	2 Q5T2I0_BRARE	Q5tzio brachydanio

32	121	11.5	763	2	Q4VB91_HUMAN	Q4vb91 homo sapien
33	121	11.5	763	2	Q4VB88_HUMAN	Q4vb88 homo sapien
34	120.5	11.4	1229	1	EGFL3_HUMAN	Q75095 homo sapien
35	120	11.4	573	2	Q5W9G8_HUMAN	Q5w9g8 homo sapien
36	120	11.4	740	2	Q6PIA2_HUMAN	Q6pia2 homo sapien
37	120	11.4	1123	2	Q75QY0_HUMAN	Q75qy0 homo sapien
38	120	11.4	4349	1	FAT2_HUMAN	Q9nyq8 homo sapien
39	119.5	11.3	1957	2	Q4SU28_TETNG	Q4su28 tetraodon n
40	119	11.3	2556	1	NOTC1_HUMAN	P46531 homo sapien
41	118.5	11.3	1722	2	Q19350_CABEL	Q19350 caenorhabdi
42	117.5	11.2	477	1	URT2_DESRO	P15638 desmodus ro
43	117	11.1	1270	2	Q9GPN0_CAEBR	Q9gpn0 caenorhabdi
44	116.5	11.1	329	2	Q7QER8_ANOGR	Q7qer8 anopheles g
45	116	11.0	737	2	Q6IR63_XENLA	Q6ir63 xenopus lae

#### ALIGNMENTS

RESULT 1  
ID TDGF1\_HUMAN STANDARD; PRT; 188 AA.  
AC P13385;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DE 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Teratocarcinoma-derived growth factor 1 precursor (Epidermal growth factor-like cripto protein CR1) (Cripto-1 growth factor) (CRGF).  
GN Name=TDGF1; Synonyms=CR1PTO;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=90005403; PubMed=2792079;  
RA Ciccodicola A., Dono R., Obici S., Zollo M., Persico M.G.;  
RT "Molecular characterization of a gene of the 'EGF family' expressed in undifferentiated human NTERA2 teratocarcinoma cells.";  
RL EMBO J. 8:1987-1991(1989).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=91353571; PubMed=1882841;  
RA Dono R., Montuori N., Rocchi M., de Ponti-Zilli L., Ciccodicola A., Persico M.G.;  
RT "Isolation and characterization of the CR1PTO autosomal gene and its X-linked related sequence.";  
RL Am. J. Hum. Genet. 49:555-565(1991).  
RN [3]  
RP PROTEIN SEQUENCE OF 31-45.  
RX PubMed=15340161; DOI=10.1110/ps.04682504;  
RA Zhang Z., Henzel W.J.;  
RT "Signal peptide prediction based on analysis of experimentally verified cleavage sites.";  
RL Protein Sci. 13:2819-2824(2004).  
RN [4]  
RP DISULFIDE BONDS.  
RX MEDLINE=22800780; PubMed=12919325;  
DOI=10.1046/j.1432-1033.2003.03749.x;  
RA Foley S.F., Van Vlijmen H.W., Boynton R.E., Adkins H.B., Cheung A.E., Singh J., Sanicola M., Young C.N., Wen D.;  
RT "The CR1PTO/PRL-1/CRYPTIC (CFC) domain of human Cripto.";  
RL Eur. J. Biochem. 270:3610-3618(2003).  
CC -!- FUNCTION: Could play a role in the determination of the epiblastic cells that subsequently give rise to the mesoderm.  
CC -!- TISSUE SPECIFICITY: Preferentially expressed in gastric and colorectal carcinomas than in their normal counterparts.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its

use as long as its content is in no way modified and this statement is not removed.

CC EMBL; M96955; AA61134.1; -; Genomic\_DNA.  
 CC EMBL; X14253; CA32467.1; -; mRNA.  
 DR PIR; B39787; A30362.  
 DR HSSP; P00740; IEDM.  
 DR Ensembl; ENSG00000163828; Homo sapiens.  
 DR HGNC; HGNC:11701; TDGF1.  
 DR H-InvDB; HIX0003250; -.  
 DR MIM; 187395; -.  
 DR GO; GO:000986; C:cell surface; IDA.  
 DR GO; GO:0019897; C:extrinsic to plasma membrane; ISS.  
 DR GO; GO:0008083; F:growth factor activity; IDA.  
 DR GO; GO:0000187; P:activation of MAPK; IDA.  
 DR GO; GO:0001763; P:branching morphogenesis; TAS.  
 DR GO; GO:0030154; P:cell differentiation; TAS.  
 DR GO; GO:0008595; P:determination of anterior/posterior axis, e. . .; ISS.  
 DR GO; GO:0009790; P:embryonic development; TAS.  
 DR GO; GO:0007507; P:heart development; IDA.  
 DR GO; GO:0030879; P:mammary gland development; TAS.  
 DR GO; GO:0043066; P:negative regulation of apoptosis; IDA.  
 DR GO; GO:0018105; P:peptidyl-serine phosphorylation; IDA.  
 DR GO; GO:0030335; P:positive regulation of cell migration; IDA.  
 DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.  
 DR GO; GO:0050731; P:positive regulation of peptidyl-tyrosine ph. . .; IDA.  
 DR GO; GO:0009666; P:regulation of signal transduction; IDA.  
 DR InterPro; IPR000742; EGF\_2.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR006210; IEGF.  
 DR Pfam; PF00008; EGF; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR PROSITE; PS00022; EGF 1; 1.  
 DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 KW Direct protein sequencing; EGF-like domain; Glycoprotein;  
 KW Growth factor; Polymorphism; Signal.  
 FT SIGNAL 1 30  
 FT CHAIN 31 188 Teratocarcinoma-derived growth factor 1.  
 FT DOMAIN 78 107 EGF-like.  
 FT CARBOHYD 79 79 N-linked (GlcNAc. . .) (Potential).  
 FT DISULFID 82 89 By similarity.  
 FT DISULFID 83 95 By similarity.  
 FT DISULFID 97 106 By similarity.  
 FT DISULFID 115 133  
 FT DISULFID 128 149  
 FT DISULFID 131 140  
 FT VARIANT 43 43 Y -> D (in dbSNP:2293025).  
 FT FTID=VAR 021903.  
 SQ SEQUENCE 188 AA; 21169 MW; AE58727D0F27D866 CRC64;

Query Match 100.0%; Score 1053; DB 1; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-89;  
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDCRWARFSYVIMAIKSKVFEGLVAGLGHQEFARPSRGYLAFRDSDIWPQEPAIR 60  
 DB 1 MDCRWARFSYVIMAIKSKVFEGLVAGLGHQEFARPSRGYLAFRDSDIWPQEPAIR 60  
 QY 61 PRSSORVPPMGIQHSKELNRTCTCLNGGTCMLGSFCACPPSFYGRNCEHDVRCNGSVPH 120  
 DB 61 PRSSORVPPMGIQHSKELNRTCTCLNGGTCMLGSFCACPPSFYGRNCEHDVRCNGSVPH 120  
 QY 121 DTWLPKCSLCKWGHQRCFPAQLPGCDGLVMDHLVASRTPELPSARTTTLMLVGI 180  
 DB 121 DTWLPKCSLCKWGHQRCFPAQLPGCDGLVMDHLVASRTPELPSARTTTLMLVGI 180  
 QY 181 CLSIQSY 188  
 DB 181 CLSIQSY 188

RESULT 2

Q8TCC1 HUMAN  
 ID Q8TCC1 HUMAN PRELIMINARY; PRT; 188 AA.  
 AC Q8TCC1  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Teratocarcinoma-derived growth factor 1.  
 GN Name=TDGF1;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE  
 RC TISSUE=Brain, and Kidney;  
 RX MEDLINE=2338257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalios D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain;  
 RG NIH MGC Project;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC022393; AAH22393.1; -; mRNA.  
 DR EMBL; BC067844; AAH67844.1; -; mRNA.  
 DR PIR; A39787; A39787.  
 DR HSSP; P00740; IEDM.  
 DR Ensembl; ENSG00000163828; Homo sapiens.  
 DR InterPro; IPR000742; EGF 2.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR006210; IEGF.  
 DR Pfam; PF00008; EGF; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 KW EGF-like domain; Transmembrane.  
 SQ SEQUENCE 188 AA; 21141 MW; E1B3A8461C855FFF CRC64;  
 Query Match 99.6%; Score 1049; DB 2; Length 188;  
 Best Local Similarity 99.5%; Pred. No. 8e-89;  
 Matches 187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MDCRWARFSYVIMAIKSKVFEGLVAGLGHQEFARPSRGYLAFRDSDIWPQEPAIR 60  
 DB 1 MDCRWARFSYVIMAIKSKVFEGLVAGLGHQEFARPSRGYLAFRDSDIWPQEPAIR 60  
 QY 61 PRSSORVPPMGIQHSKELNRTCTCLNGGTCMLGSFCACPPSFYGRNCEHDVRCNGSVPH 120  
 DB 61 PRSSORVPPMGIQHSKELNRTCTCLNGGTCMLGSFCACPPSFYGRNCEHDVRCNGSVPH 120



```

QY 66 RVPFMGIQHSKELNRTCCCLNGTCTMLGSCFACPPSFYGRNCEHVRKENCSCVPHDWTLP 125
DB 61 FVANSNGIQHSKELNRTCCCLNGTCTMLGSCFACPPSFYGRNCEHVRKENCSCVPHDWTLP 120
QY 126 KKSCLCKCWHGOLRCFPOAFIPGCDGLVMDHVLASRTPELPSPARTTTFMVLGVCLSIQ 185
DB 121 RKCSMCKCWHGOLRCFPOAFIPGCDGLVMDHVLASRTPELPSPARTTTFMVLGVCLSIQ 178
QY 186 SYY 188
DB 179 SYY 181

RESULT 5
TDGFI_MOUSE STANDARD; PRT; 171 AA.
AC P51865;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Teratocarcinoma-derived growth factor precursor (Epidermal growth
DE factor-like Cripto protein) (Cripto growth factor).
GN Name=tdgf1; Synonyms=Cripto;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Embryo;
RX MEDLINE=94094736; PubMed=7916676;
RA Liguori G., Tucci M., Montuori N., Dono R., Pacifico A.F.,
RA Persico M.G.;
RA "Characterization of the mouse Tdgf1 gene and Tdgf pseudogenes.";
RL Mamm. Genome 7:344-348(1996).
CC -!- FUNCTION: Could play a role in the determination of the epiblastic
CC cells that subsequently give rise to the mesoderm.
CC -!- TISSUE SPECIFICITY: Expressed at low level in specific organs of
CC the adult animal such as spleen, heart, lung and brain. During
CC gastrulation, expressed in the forming mesoderm. In later stages
CC of the developing heart, expression is restricted to the truncus
CC arteriosus.
CC -!- DEVELOPMENTAL STAGE: First expressed prior to the onset of
CC gastrulation (early streak stage), then continues throughout
CC embryonic development.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; M87321; AAA37459.1; -; mRNA.
CC EMBL; X94083; CAA63827.1; -; Genomic_DNA.
CC PIR; I49612; I49612.
CC HSSP; P00740; 1EDM.
CC
CC Ensembl; ENSMUSG00000032494; Mus musculus.
CC MGI; MGI:98659; tdgf1.
CC
CC GO; GO:0005615; C:extracellular space; TAS.
CC GO; GO:0019897; C:extrinsic to plasma membrane; IDA.
CC GO; GO:0001763; P:branching morphogenesis; IDA.
CC GO; GO:0030154; P:cell differentiation; TAS.

```

```

DR GO; GO:0008595; P:determination of anterior/posterior axis, e. . .; IMP.
DR GO; GO:0007507; P:heart development; IDA.
DR GO; GO:0030335; P:positive regulation of cell migration; IDA.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00008; EGF_1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00026; EGF_3; 1.
KW EGF-like domain; Glycoprotein; Growth factor; Signal.
FT SIGNAL 1 17 Potential.
FT CHAIN 18 171 Teratocarcinoma-derived growth factor.
FT DOMAIN 62 91 EGF-like.
FT CARBOHYD 63 63 N-linked (GlcNAc. .) (Potential).
FT DISULFID 66 73 By similarity.
FT DISULFID 67 79 By similarity.
FT DISULFID 81 90 By similarity.
FT DISULFID 99 117 By similarity.
FT DISULFID 112 133 By similarity.
FT DISULFID 115 124 By similarity.
SQ SEQUENCE 171 AA; 18754 MW; C52051AEACDBS380 CRC64;

Query Match 61.4%; Score 647; DB 1; Length 171;
Best Local Similarity 67.8%; Pred. No. 9.9e-52;
Matches 120; Conservative 14; Mismatches 31; Indels 12; Gaps 2;

QY 6 MARPSYSVIWIMAIKVPFELGLVAGLGHQEFARPSRGVLAFRDSDSIWQPEEPAIRPSQ 65
DB 1 MGYFSSSVLLVAISSAFEGFVAGRD-----LAIRDNSIWQKEFAVDRSRQ 49
[1]
QY 66 RVPFMGIQHSKELNRTCCCLNGTCTMLGSCFACPPSFYGRNCEHVRKENCSCVPHDWTLP 125
DB 50 FVPSVGIQNSKLNKTCCLNGTCTMLGSCFACPPSFYGRNCEHVRKENCSCVPHDWTLP 109
[1]
QY 126 KKSCLCKCWHGOLRCFPOAFIPGCDGLVMDHVLASRTPELPSPARTTTFMVLGVCL 182
DB 110 KKSCLRCWHGOLHCLPQTFLPGCDGHVMDQDLKASRTPCQTPSV-TTTFMLAGACL 165

RESULT 6
QYTDGFI_MOUSE PRELIMINARY; PRT; 171 AA.
AC Q7T006;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Tdgf1 protein.
GN Name=tdgf1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Blasocyst;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.I., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailis D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

```



"Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]

## NUCLEOTIDE SEQUENCE

RP STRAIN=CS7BL/6J; TISSUE=Blastocyst;  
RA Strausberg R.;  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC052646; AAH52646.1; -; mRNA.  
DR HSSP; P00740; 1EDM.  
DR MGI; MGI:198658; Tdgfl.  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR GO; GO:0019897; C:extrinsic to plasma membrane; IDA.  
DR GO; GO:0008595; P:determination of anterior/posterior axis, e. . .; IMP.  
DR InterPro; IPR00742; EGF 2.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR006210; IEGF.  
DR Pfam; PF00008; EGF; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN 1.  
DR PROSITE; PS00026; EGF\_3; 1.  
SQ SEQUENCE 171 AA; 18654 MW; C53400EBACDB5380 CRC64;

Query Match 60.8%; Score 640; DB 2; Length 171;  
Best Local Similarity 67.2%; Pred. No. 4.4e-51;  
Matches 119; Conservative 14; Mismatches 32; Indels 12; Gaps 2;

QY 6 MARFSYVIMWATSKVFLGLVAGLQHFARPSRGVLAFRDSDIWPQEPAIRPRSSQ 65  
DB 1 MGYFSSVVLVAISSAFEGFVAGRD-----LAIRDSIWDQREPAVDRSFQ 49  
QY 66 RVPPMGIQHSKELNRTCLNGGTCLMGSCFACPPSFYGRNCEHDVRCNGSVPHDTWLP 125  
DB 50 FVPSVGIQNSKSLNKTCLNGGTCLMGSCFACPPSFYGRNCEHDVRCNGSVILHGTWLP 109  
QY 126 KKSCLCKWGHLCRCFQAFLPGCDGLVDEHLVASFTPELPSPARTTMTLVGICL 182  
DB 110 KKSCLCKWGHLCRCFQAFLPGCDGLVDEHLVASFTPELPSPARTTMTLVGICL 165

## RESULT 7

Q91A22\_BRARE  
ID Q91A22\_BRARE PRELIMINARY; PRT; 66 AA.  
AC Q91A22;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Crypto (Fragment).  
GN Name=crypto;  
OS Brachydanio rerio (zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.

RA Del Giacco L., Cotelli F.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF151352; AAF73188.1; -; mRNA.  
DR ZFIN; ZDB-GENE-000710-5; crypto.  
FT NON TER 1  
FT NON TER 66  
SQ SEQUENCE 66 AA; 7275 MW; 5C6ABBI6D5832871 CRC64;

Query Match 34.5%; Score 363; DB 2; Length 66;  
Best Local Similarity 92.4%; Pred. No. 6.7e-26;  
Matches 61; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 77 ELNRTCLNGGTCLMGSCFACPPSFYGRNCEHDVRCNGSVPHDTWLPKKCSI-CKWHG 136  
DB 1 ELNRTCLNGGTCLMGSCFACPPSFYGRNCEHDVRCNGSVPHDTWLPKKCSI-CKWHG 60  
QY 137 QLRCPF 142

Db 61 QLRCPF 66  
|||||

## RESULT 8

OS7517\_BRARE  
ID OS7517\_BRARE PRELIMINARY; PRT; 183 AA.  
AC OS7517;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE One-eyed pinhead short form protein.  
GN Name=oeep;  
OS Brachydanio rerio (zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.

RA MEDLINE=98117252; PubMed=9458048; DOI=10.1016/S0092-8674(00)80918-6;  
RA Zhang J., Talbot W.S., Schier A.F.;  
RT "Positional cloning identifies zebrafish one-eyed pinhead as a  
RT permissive EGF-related ligand required during gastrulation.";  
RL Cell 92:241-251(1998).  
DR EMBL; AF041441; AAC04339.1; -; mRNA.  
DR HSSP; P00750; ITPG.  
DR Ensembl; ENSDARG0000035095; Danio rerio.  
DR ZFIN; ZDB-GENE-990415-198; oep.  
DR InterPro; IPR006209; EGF like.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN 1.  
SQ SEQUENCE 183 AA; 20404 MW; 1471447894E1851B CRC64;

Query Match 26.1%; Score 274.5; DB 2; Length 183;

Best Local Similarity 47.4%; Pred. No. 2.7e-17;

Matches 46; Conservative 18; Mismatches 32; Indels 1; Gaps 1;

QY 55 EEPAIRPRSSQ-RVPPMGIQHSKELNRTCLNGGTCLMGSCFACPPSFYGRNCEHDVRC 113  
DB 55 QTPQRHNAEALPFVGLTGVAKQSRCTCKNGGTCLMGSCFACPKVFTGRSCYDRLR 114

QY 114 NCGSVPHDTWLPKKCSLCKWGHLCRCFQAFLPGCD 150

DB 115 DCGVPHGEVWQKSCYRCGCGYLLHCFPHVFSKDCD 151

## RESULT 9

Q50415\_BRARE  
ID Q50415\_BRARE PRELIMINARY; PRT; 183 AA.  
AC Q50415;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Oep protein.  
GN Name=oeep;  
OS Brachydanio rerio (zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.

RA TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci F., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC095007; AAH95007.1; -; mRNA.
SQ SEQUENCE 183 AA; 20374 MW; 5571447894E18507 CRC64;

  Query Match      26.1%; Score 274.5; DB 2; Length 183;
  Best Local Similarity 47.4%; Pred. No. 2.7e-17;
  Matches 46; Conservative 18; Mismatches 32; Indels 1; Gaps 1;

QY 55 BEPAIRPSSQ-RVPPMGIOHSEKLNRTCCLLNGGTCMLGSCFACPPSFYGRNCEHDVRKE 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 55 QTPQQRHNAEALPFVGLTGVAQKSRCTCKNGGTCILGSCFACCPKYTGRSCEYDERLR 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 114 NCGSVPHDTWLPKCKSLCKCMWHGQRLRCFPQAFPLPGCD 150
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 115 DCGVIPGHEWQKGCYCRGCGYGLLHCFPHVFSKDCD 151
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
ID O57516 BRARE PRELIMINARY; PRT; 190 AA.
AC O57516;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE One-eyed pinhead long form protein.
GN Name=oeop;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98117252; PubMed=9458048; DOI=10.1016/S0092-8674(00)80918-6;
RA Zhang J., Talbot W.S., Schier A.F.;
RT "Positional cloning identifies zebrafish one-eyed pinhead as a
RT permissive EGF-related ligand required during gastrulation.";
RL Cell 92:241-251(1998).
DR EMBL; AF041440; AAC04338.1; -; mRNA.
DR HSSP; P00750; 1TPG.
DR Ensembl; ENSDARG0000035095; Danio rerio.
DR ZFIN; ZDB-GENE-990415-198; oep.
DR InterPro; IPR006209; EGF like.
DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
SQ SEQUENCE 190 AA; 21199 MW; 8D127CAD0625D28A CRC64;

  Query Match      26.1%; Score 274.5; DB 2; Length 190;
  Best Local Similarity 47.4%; Pred. No. 2.8e-17;
  Matches 46; Conservative 18; Mismatches 32; Indels 1; Gaps 1;

QY 55 BEPAIRPSSQ-RVPPMGIOHSEKLNRTCCLLNGGTCMLGSCFACPPSFYGRNCEHDVRKE 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 55 QTPQQRHNAEALPFVGLTGVAQKSRCTCKNGGTCILGSCFACCPKYTGRSCEYDERLR 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 114 NCGSVPHDTWLPKCKSLCKCMWHGQRLRCFPQAFPLPGCD 150
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 115 DCGVIPGHEWQKGCYCRGCGYGLLHCFPHVFSKDCD 151
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 11
Q91649_XENLA PRELIMINARY; PRT; 190 AA.
ID Q91649_XENLA PRELIMINARY; PRT; 190 AA.
AC Q91649;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Fibroblast growth factor receptor ligand 1.
DE Name=FR1;
GN Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96069863; PubMed=7585965; DOI=10.1016/0092-8674(95)90102-7;
RA Kinoshita N., Minshull J., Kirschner M.W.;
RT "The identification of two novel ligands of the FGF receptor by a
RT yeast screening method and their activity in Xenopus development.";
RL Cell 83:621-630(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=11024280; DOI=10.1016/S0378-1119(00)00337-1;
RA Colas J.F., Schoenwolf G.C.;
RT "Subtractive hybridization identifies chick-cripto, a novel EGF-CFC
RT ortholog expressed during gastrulation, neurulation and early
RT cardiogenesis.";
RL Gene 255:205-217(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Colas J.-F., Schoenwolf G.C.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U04057; AAA83569.1; -; mRNA.
DR EMBL; AF234786; AAF78013.1; -; Genomic_DNA.
DR HSSP; P00749; 1URK.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR006209; EGF like.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
SQ SEQUENCE 190 AA; 21462 MW; 6FC1660D2B953176 CRC64;

  Query Match      26.1%; Score 274.5; DB 2; Length 190;
  Best Local Similarity 43.0%; Pred. No. 2.8e-17;
  Matches 52; Conservative 20; Mismatches 44; Indels 5; Gaps 3;

QY 61 PRSQRPVPMGIQHSKLNRTCCLLNGGTCMLGSCFACPPSFYGRNCEHDVRKNCGVPH 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 PKSTKTLFLGITDSKLNKCCQNGGTCFLGTFCICPKQFTGRHCEHRRPASCQGVPH 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 DTWLPKCKSLCKCMWHGQRLRCF-PQAFPLGCDGLVNDHLVASRTPELPSPARTTFFMLVG 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 GDWIRQGCLLCRCVSVGLHCFKPES--EDCD--VVHEKNMRSGVPRMQLSLIYCFLTAN 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 180 I 180
Db 178 L 178

RESULT 12
Q800J2_XENLA PRELIMINARY; PRT; 190 AA.
ID Q800J2_XENLA PRELIMINARY; PRT; 190 AA.
AC Q800J2;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE FR1-1.
GN Name=FR1-1;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.

```



OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20517351; PubMed=11062482; DOI=10.1038/81695;  
RA Bamford R.N., Roessler E., Burdine R., Saplakoglu U., Cruz J.D.,  
RA Splitt M., Towbin J., Marino B., Schier A.F., Shen M.M., Muenke M.,  
RA Casey B.;  
RT "Loss-of-function mutations in the EGF-CFC gene CFC1 are associated  
RT with human left-right laterality defects.";  
RL Nat. Genet. 26:365-369(2000).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Lung, and PCR rescued clones;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ushin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smaluk D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=PCR rescued clones;  
RA Director MGC Project;  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Lung;  
RA Director MGC Project;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Lung;  
RG NIH MGC Project;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF312925; AAG2475.1; -; Genomic\_DNA.  
DR EMBL; AF312769; AAG30294.1; -; mRNA.  
DR EMBL; BC069508; AAH63508.1; -; mRNA.  
DR EMBL; BC074825; AAH74825.1; -; mRNA.  
DR EMBL; BC074826; AAH74826.1; -; mRNA.  
DR HSSP; P00750; 1TPG.  
DR HGNC; HGNC:18292; CFC1.  
DR GO; GO:0007368; P.determination of left/right symmetry; NAS.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR006210; IEGF.  
DR Pfam; PF00008; EGF; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS00026; EGF\_3; 1.  
KW EGF-like domain; Transmembrane.  
SQ SEQUENCE 223 AA; 24612 MW; B52852A00ABCFA13 CRC64;  
Query Match 22.1%; Score 233; DB 2; Length 223;  
Best Local Similarity 39.1%; Pred. No. 2.2e-13;

Matches 43; Conservative 7; Mismatches 30; Indels 30; Gaps 2;  
Qy 52 WQDEP-----AIRSRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSSFCACPPS 100  
Db 68 WGPBEPLPYSAFGEASARPR-----CCRMGGTCVLGSCVCPAH 108  
Qy 101 FYGRNCEHDVRKENCQSVPHDTWLPKCKSLCKCWHGQLRCFPQAFPLPGCD 150  
Db 109 FTGRYCEHDQRRSECGALEHGAWTLRACHLCRCIFGALHCLPLQTPDRC 158

Search completed: February 3, 2006, 18:29:57  
Job time : 183.316 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 18:15:25 ; Search time 176.72 Seconds  
(without alignments)  
467.424 Million cell updates/sec

Title: US-10-693-538-1  
Perfect score: 1053  
Sequence: 1 MDCRKNARFYSYVIMWAIS.....SARTTTPMLVGICLSIQSY 188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	188	2	AAR22548 Human CRI
2	1053	100.0	188	2	AAW87630 Human CRI
3	1053	100.0	188	5	ABB77101 Human CRI
4	1053	100.0	188	5	AAO14638 Human CRI
5	1053	100.0	188	5	AAO14636 Human CRI
6	1053	100.0	188	6	ABP97176 Tumour-as
7	1053	100.0	188	6	ABP58131 Human CRI
8	1053	100.0	188	7	ADC78799 Human PRO
9	1053	100.0	188	7	ADD26058 CR-1 crip
10	1053	100.0	188	8	ADI82176 Human ter
11	1053	100.0	188	8	ADR70486 Human ter
12	1053	100.0	188	8	ADT79200 Human CRI
13	1053	100.0	188	8	ADT79197 Human CRI
14	1053	100.0	188	9	ADY85961 Human CRI
15	1053	100.0	188	9	ADZ42238 Human CRI
16	1049	99.6	188	5	AAO14727 Human var
17	1049	99.6	188	8	ADT79290 Human CRI
18	1048	99.5	188	5	ABB77109 Human CRI
19	1046	99.3	188	2	AAW29735 Homo sapi
20	1035	98.3	188	2	AAW32108 Human CRI
21	1017	96.6	188	8	ADS88697 Amino aci
22	1013	96.2	188	2	AAR22547 Protein e
23	1013	96.2	188	2	AAW25667 Protein e
24	1013	96.2	188	2	AAW19980 Human CRI

25	1013	96.2	188	2	AAW87631	AAW87631 Human CRI
26	1013	96.2	188	5	ABB77102	ABB77102 Human CRI
27	1013	96.2	188	5	AAO14637	AAO14637 Human CRI
28	1013	96.2	188	6	ABP58132	ABP58132 Human CRI
29	1013	96.2	188	7	ADC78843	ADC78843 Human PRO
30	1013	96.2	188	7	ADD26059	ADD26059 CR-2 crip
31	1013	96.2	188	8	ADT79198	ADT79198 Human CRI
32	1008	95.7	188	5	ABB77122	ABB77122 Human CRI
33	1002.5	95.2	184	8	ADO28629	ADO28629 Human CRI
34	980	93.1	183	8	ADS88699	ADS88699 Amino aci
35	975	92.6	173	8	ADS88698	ADS88698 Amino aci
36	955	90.7	174	2	AAR13326	AAR13326 Recombina
37	955	90.7	174	2	AAW32107	AAW32107 Recombina
38	955	90.7	174	8	ADO05065	ADO05065 Human CRI
39	848.5	80.6	187	6	AAE36467	AAE36467 Human col
40	805	76.4	367	5	ABB77106	ABB77106 Human CRI
41	803	76.3	339	5	ABB77103	ABB77103 Human CRI
42	800	76.0	367	5	ABB77113	ABB77113 Human CRI
43	798	75.8	339	5	ABB77110	ABB77110 Human CRI
44	777	73.8	367	5	ABB77119	ABB77119 Human CRI
45	775	73.6	339	5	ABB77123	ABB77123 Human CRI

## ALIGNMENTS

## RESULT 1

AAR22548	
ID	AAR22548 standard; protein; 188 AA.
XX	
XX	AAR22548;
XX	
DT	25-MAR-2003 (revised)
DT	17-DEC-2001 (revised)
DT	28-AUG-1992 (first entry)
XX	
DE	Human CRIPTO CR-1 protein.
XX	
KW	cell proliferation; tumour; CR-3; transforming growth factor;
KW	epidermal growth factor.
XX	
OS	Homo sapiens.
XX	
PN	USN7749001-N.
XX	
PD	25-FEB-1992.
XX	
PF	23-AUG-1991; 91US-00749001.
XX	
PR	23-AUG-1991; 91US-00749001.
XX	
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX	
PI	Salomon D, Persico M;
XX	
DR	WPI; 1992-123675/15.
DR	N-PSDB; AAQ22495.
XX	
PT	New crpto gene CR-1 and crpto-related gene CR-3 genomic DNA - CR-3
PT	protein and anti-CR-3 antibodies, useful in immunoassay to detect CR-3 as
PT	tumour specific marker.
XX	
PS	Disclosure; Page 22; 44pp; English.
XX	
CC	The CR-1 genomic clone which codes for this protein can be used in
CC	transgenic animals to examine the effects of overexpression of this gene
CC	on development and tumorigenicity and to study regulation of CR-1 gene.
CC	See also AAQ22494. (Note: Revised entry submitted to correct the patent
CC	number format of US Government-owned NTIS applications to prevent clashes
CC	with ongoing US granted patent numbers. For further information please
CC	visit the Derwent web site at www.derwent.com/dwpi/updates/ntis_us.html.)
CC	(Updated on 25-MAR-2003 to correct PA field.)
XX	

SQ Sequence 188 AA;

Query Match 100.0%; Score 1053; DB 2; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-82;  
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCRMAFYSYVIWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEPAPR 60  
 |||||  
 DB 1 MDCRMAFYSYVIWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEPAPR 60  
 |||||

QY 61 PRSSQRPVPMGIQHSKELNRTCCNGGTCMLGSCFACPPSFYGRNCEHDVRKENC SVPH 120  
 |||||  
 DB 61 PRSSQRPVPMGIQHSKELNRTCCNGGTCMLGSCFACPPSFYGRNCEHDVRKENC SVPH 120  
 |||||

QY 121 DTWLPKKSLCKCWGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180  
 |||||  
 DB 121 DTWLPKKSLCKCWGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180  
 |||||

QY 181 CLSIQSY 188  
 |||||  
 DB 181 CLSIQSY 188  
 |||||

RESULT 2  
 AA087630  
 ID AA087630 standard; protein; 188 AA.  
 XX  
 AC AA087630;  
 XX  
 DT 03-MAR-1999 (first entry)  
 XX  
 DE Human CRIPTO-related (CR-1) protein.  
 XX  
 KW human CRIPTO-related protein; CR-1; antibody; CR-3; proliferation;  
 KW differentiation; transformation; mesenchymal cell; epithelial cell;  
 KW tumour specific marker; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5854399-A.  
 XX  
 PD 29-DEC-1998.  
 XX  
 PF 05-JUN-1995; 95US-00464023.  
 XX  
 PR 23-AUG-1991; 91US-00749001.  
 PR 17-NOV-1993; 93US-00154198.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Persica MG, Salomon DS;  
 XX  
 DR WPI; 1999-095001/08.  
 DR N-PSDB; AA083923.  
 XX  
 PT New antibody to human CRIPTO-related polypeptide-3 - used for the  
 PT detection of CRIPTO-related polypeptide-3 which is involved in the  
 PT regulation of the proliferation, differentiation and transformation of  
 PT cells.  
 XX  
 PS Example 2; Col 17-18; 26pp; English.  
 XX  
 CC The present sequence represents a human CRIPTO-related protein CR-1. The  
 CC specification describes an antibody which has binding affinity to CR-3  
 CC and not to CR-1. The antibodies can be used for the detection of CR-3. CR  
 CC -3 is a regulatory molecule involved in regulating the proliferation,  
 CC differentiation, and transformation of various mesenchymal and epithelial  
 CC cells. In addition expression of CR-3 may serve as a tumour specific  
 CC marker that may have applicability in the diagnosis, prognosis and  
 CC possible treatment of specific types of cancer  
 XX  
 SQ Sequence 188 AA;

Query Match 100.0%; Score 1053; DB 2; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-82;  
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCRMAFYSYVIWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEPAPR 60  
 |||||  
 DB 1 MDCRMAFYSYVIWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEPAPR 60  
 |||||

QY 61 PRSSQRPVPMGIQHSKELNRTCCNGGTCMLGSCFACPPSFYGRNCEHDVRKENC SVPH 120  
 |||||  
 DB 61 PRSSQRPVPMGIQHSKELNRTCCNGGTCMLGSCFACPPSFYGRNCEHDVRKENC SVPH 120  
 |||||

QY 121 DTWLPKKSLCKCWGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180  
 |||||  
 DB 121 DTWLPKKSLCKCWGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180  
 |||||

QY 181 CLSIQSY 188  
 |||||  
 DB 181 CLSIQSY 188  
 |||||

RESULT 3  
 ABB77101  
 ID ABB77101 standard; protein; 188 AA.  
 XX  
 AC ABB77101;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Human CRIPTO-1 full length protein.  
 XX  
 KW Human; CRIPTO-1; CR-1; mutant; tumour; cell proliferation.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..30  
 FT /label= Signal\_peptide  
 FT Protein 31..188  
 FT /label= Mature\_Cripto-1  
 XX  
 PN WO200222808-A2.  
 XX  
 PD 21-MAR-2002.  
 XX  
 PF 18-SEP-2001; 2001WO-US029066.  
 XX  
 PR 18-SEP-2000; 2000US-0233148P.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI Williams KP, Foley S, Schiffer S, Domon B, Sanicola-Nadel M;  
 XX  
 DR WPI; 2002-339868/37.  
 DR N-PSDB; ABL55852.  
 XX  
 PT New mutant form of CRIPTO (teratocarcinoma-derived growth factor), useful  
 PT for treating cell proliferation, especially cancer, comprises amino acid  
 PT change that prevents fucosylation at Thr88.  
 XX  
 PS Claim 2; Fig 1; 41pp; English.  
 XX  
 CC The sequence represents the full length human Crip1-1 protein. The  
 CC invention relates to a novel mutant Crip1-1 polypeptide, or its functional  
 CC fragment, having at least one amino acid alteration at positions 86, 87  
 CC or 88. The mutant polypeptide, or its chimera, is used to inhibit growth  
 CC of tumour cells, in vivo or in vitro, particularly for treating breast,  
 CC ovarian, renal, colorectal, uterine, prostatic, lung, bladder or central  
 CC nervous system cancers, melanoma and leukaemia, also generally for  
 CC treating undesired cell proliferation  
 XX  
 SQ Sequence 188 AA;

```
Query Match      100.0%; Score 1053; DB 5; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-82;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCRKFARFYSYVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEEPAIR 60
    |||||
DB 1 MDCRKFARFYSYVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEEPAIR 60
    |||||

QY 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCHEHDVRKENCSSVPH 120
    |||||
DB 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCHEHDVRKENCSSVPH 120
    |||||

QY 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCHEHDVRKENCSSVPH 120
    |||||
DB 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCHEHDVRKENCSSVPH 120
    |||||

QY 121 DTWLPKKCSLCKWHGQLRCFPOAFLPGCDGLVNDHLVASRTPPELPPSARTTTTFLVLGI 180
    |||||
DB 121 DTWLPKKCSLCKWHGQLRCFPOAFLPGCDGLVNDHLVASRTPPELPPSARTTTTFLVLGI 180
    |||||

QY 181 CLSIQSYI 188
    |||||
DB 181 CLSIQSYI 188
    |||||

RESULT 4
AAO14638
ID AAO14638 standard; protein; 188 AA.
XX
AC AAO14638;
XX
XX
DT 13-JUN-2002 (first entry)
XX
DE Human cripto protein.
XX
KW Human; cripto protein; epidermal growth factor family; EGF family;
KW gene therapy; protein therapy; vaccine; lung cancer; colon cancer;
KW colorectal cancer; breast cancer.
XX
OS Homo sapiens.
XX
PN WO200216413-A2.
XX
PD 28-FEB-2002.
XX
PF 20-AUG-2001; 2001WO-EP009646.
XX
PR 24-AUG-2000; 2000GB-00020953.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cassart J, Coche T, Palmantier RM, Vinals Y De BassolsC;
XX WPI; 2002-280910/32.
XX DR N-PSDB; AAL42198.
XX
PT New Cripto tumor polynucleotides, useful in medicine, particularly for
PT treating tumors (e.g. lung, colorectal, colon or breast tumor) that
PT express a cripto antigen.
XX
PS Diagnostic; Page 95; 134pp; English.
XX
CC The invention comprises the nucleotide and amino acid sequences of human
CC cripto proteins. Cripto is a 188 amino acid protein that shares
CC homologues with the epidermal growth factor (EGF) family. The cripto
CC nucleic acid and protein sequences of the invention are useful for
CC treating (e.g. gene therapy, protein therapy, vaccine) cancer, such as:
CC lung, colon, colorectal or breast. The present amino acid sequence
CC represents a human cripto protein
XX
SQ Sequence 188 AA;

Query Match      100.0%; Score 1053; DB 5; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-82;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCRKFARFYSYVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEEPAIR 60
    |||||
DB 1 MDCRKFARFYSYVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEEPAIR 60
    |||||

QY 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCHEHDVRKENCSSVPH 120
    |||||
DB 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCHEHDVRKENCSSVPH 120
    |||||

QY 121 DTWLPKKCSLCKWHGQLRCFPOAFLPGCDGLVNDHLVASRTPPELPPSARTTTTFLVLGI 180
    |||||
DB 121 DTWLPKKCSLCKWHGQLRCFPOAFLPGCDGLVNDHLVASRTPPELPPSARTTTTFLVLGI 180
    |||||

QY 181 CLSIQSYI 188
    |||||
DB 181 CLSIQSYI 188
    |||||

RESULT 5
AAO14636
ID AAO14636 standard; protein; 188 AA.
XX
AC AAO14636;
XX
XX
DT 13-JUN-2002 (first entry)
XX
DE Human cripto 1 protein.
XX
KW Human; cripto 1 protein; epidermal growth factor family; EGF family;
KW gene therapy; protein therapy; vaccine; lung cancer; colon cancer;
KW colorectal cancer; breast cancer.
XX
OS Homo sapiens.
XX
PN WO200216413-A2.
XX
PD 28-FEB-2002.
XX
PF 20-AUG-2001; 2001WO-EP009646.
XX
PR 24-AUG-2000; 2000GB-00020953.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cassart J, Coche T, Palmantier RM, Vinals Y De BassolsC;
XX WPI; 2002-280910/32.
XX DR N-PSDB; AAL42198.
XX
PT New Cripto tumor polynucleotides, useful in medicine, particularly for
PT treating tumors (e.g. lung, colorectal, colon or breast tumor) that
PT express a cripto antigen.
XX
PS Claim 1; Page 94; 134pp; English.
XX
CC The invention comprises the nucleotide and amino acid sequences of human
CC cripto proteins. Cripto is a 188 amino acid protein that shares
CC homologues with the epidermal growth factor (EGF) family. The cripto
CC nucleic acid and protein sequences of the invention are useful for
CC treating (e.g. gene therapy, protein therapy, vaccine) cancer, such as:
CC lung, colon, colorectal or breast. The present amino acid sequence
CC represents the human cripto 1 protein
XX
SQ Sequence 188 AA;

Query Match      100.0%; Score 1053; DB 5; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e-82;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCRKFARFYSYVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEEPAIR 60
    |||||
DB 1 MDCRKFARFYSYVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDSDSIWPOEEPAIR 60
    |||||

QY 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCHEHDVRKENCSSVPH 120
    |||||
```

Db 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCLMGSCFACACPPSYGRNCEHDVVRKENGCSVPH 120  
QY 121 DTWLPKCKSLCKCWGQLRCFPOAFLPGCDGLVMDHLVASRTPELPSPARTTTFMVLGI 180  
Db 121 DTWLPKCKSLCKCWGQLRCFPOAFLPGCDGLVMDHLVASRTPELPSPARTTTFMVLGI 180  
QY 181 CLSIQSY 188  
Db 181 CLSIQSY 188  
RESULT 6  
ID ABP97176 standard; protein; 188 AA.  
AC ABP97176;  
DT 01-JUL-2003 (first entry)  
DE Tumour-associated antigenic target protein TAT177 SEQ ID NO:58.  
XX Human; tumour-associated antigenic target; TAT; tumour; diagnosis;  
KW cancer.  
OS Homo sapiens.  
XX WO2003024392-A2.  
PN 27-MAR-2003.  
PD 11-SEP-2002; 2002WO-US028859.  
PF 18-SEP-2001; 2001US-0323268P.  
PR 19-OCT-2001; 2001US-0339227P.  
PR 07-NOV-2001; 2001US-0336827P.  
PR 20-NOV-2001; 2001US-0331906P.  
PR 02-JAN-2002; 2002US-0345444P.  
PR 03-APR-2002; 2002US-0369724P.  
PR 19-AUG-2002; 2002US-0404809P.  
XX (GETH ) GENENTECH INC.  
PA Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;  
PI Williams PM, Wu TD, Zhang Z;  
XX WPI; 2003-354551/33.  
DR N-PSDB; ACC49494.  
XX New antibodies against tumor-associated antigenic target polypeptide,  
PT useful for treating or diagnosing tumors or cancers in mammals, e.g.  
PT prostate cancer, lung cancer, prostate adenocarcinomas or renal cell  
PT carcinomas.  
PS Claim 2; Fig 58; 285pp; English.  
XX ACC49493 to ACC49552 encode the human tumour-associated antigenic target  
CC (TAT) proteins given in ABP97175 to ABP97234. The present invention  
CC describes an isolated antibody that binds to a polypeptide having at  
CC least 80 % sequence identity to any of the 60 150-800 residue amino acid  
CC sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking  
CC its associated signal peptide, encoded by any of the 60 2000-3000 base  
CC pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have  
CC cytostatic activity. The antibody can be used for treating or diagnosing  
CC tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast  
CC cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal  
CC cell carcinomas, or thyroid cancer  
XX Sequence 188 AA;  
SQ Query Match 100.0%; Score 1053; DB 6; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.3e-82;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCRKMARFSYVIWMAISKVFLVAGLGHQEFARPSRGYLAFRDDSIWPQEEPAIR 60  
Db 1 MDCRKMARFSYVIWMAISKVFLVAGLGHQEFARPSRGYLAFRDDSIWPQEEPAIR 60  
QY 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCLMGSCFACACPPSYGRNCEHDVVRKENGCSVPH 120  
Db 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCLMGSCFACACPPSYGRNCEHDVVRKENGCSVPH 120  
QY 121 DTWLPKCKSLCKCWGQLRCFPOAFLPGCDGLVMDHLVASRTPELPSPARTTTFMVLGI 180  
Db 121 DTWLPKCKSLCKCWGQLRCFPOAFLPGCDGLVMDHLVASRTPELPSPARTTTFMVLGI 180  
QY 181 CLSIQSY 188  
Db 181 CLSIQSY 188  
RESULT 7  
ID ABP58131 standard; protein; 188 AA.  
XX ABP58131;  
AC ABP58131;  
DT 28-MAR-2003 (first entry)  
DE Human Cripto CR-1.  
XX Cripto; human; antitumour; tumour.  
KW Cripto; human; antitumour; tumour.  
XX Homo sapiens.  
OS Location/Qualifiers  
FH 31..188  
FT Domain /note= "extracellular domain"  
FT Region 46..62  
FT Domain /note= "epitope region specifically described in Claim 8"  
FT Domain 75..150  
FT Domain /note= "ligand/receptor binding domain"  
FT Domain 75..112  
FT Domain /note= "EGF-like domain"  
FT Domain 114..150  
FT Domain /note= "Cysteine-rich domain"  
FT Domain 169..188  
FT Domain /note= "transmembrane domain"  
XX WO200288170-A2.  
PN 07-NOV-2002.  
XX 17-APR-2002; 2002WO-US011950.  
XX 26-APR-2001; 2001US-0286782P.  
PR 17-MAY-2001; 2001US-0293020P.  
PR 26-JUN-2001; 2001US-0301091P.  
PR 22-MAR-2002; 2002US-0367002P.  
XX (BIOJ ) BIOGEN INC.  
PA Sanicola-Nadel M, Williams K, Schiffer S, Rayhorn P;  
PI WPI; 2003-120460/11.  
XX New Cripto-binding antibodies, useful for modulating the growth of tumor  
PT cells, or for treating a subject having e.g. breast, testicular, colon or  
PT lung tumor, or suffering from a disorder associated with abnormal levels  
PT of Cripto.  
XX Claim 2; Page 52-53; 54pp; English.  
PS The present sequence is the amino acid sequence of human Cripto CR-1, a  
XX cell surface protein whose over-expression is associated with many tumour  
CC types. The invention provides antibodies that specifically bind to Cripto  
CC CR-1 or CR-3 (see ABP58132) and which are useful for modulating Cripto



CC signalling or protein interaction, and/or block the interaction between  
 CC Cripto and ALK4, and/or modulate the growth of tumour cells. The  
 CC antibodies specifically bind an epitope in the ligand/receptor binding  
 CC domain of Cripto, especially in an EGF-like domain, or a Cys-rich domain,  
 CC or a domain spanning amino acid residues 46-62 of Cripto. A claimed  
 CC composition for administration to a subject having a tumour that  
 CC expresses Cripto comprises at least one of the antibodies and may also  
 CC include a chemotherapeutic agent. The composition can be used to modulate  
 CC the growth of tumour cells in vitro or in vivo, and is used in claimed  
 CC methods of treating a subject having a tumour which over-expresses Cripto  
 CC or a condition associated with undesired cell proliferation. The tumour  
 CC is selected from breast, testicular, colon, lung, ovary, bladder,  
 CC uterine, cervical, pancreatic and stomach tumour. Also claimed are  
 CC immunoassay methods of determining whether a tissue expresses Cripto or  
 CC whether a cell line over-expresses Cripto  
 XX  
 XX

Query Match 100.0%; Score 1053; DB 6; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-82;  
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCRKMARFSYSVITWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
 |||||

Db 1 MDCRKMARFSYSVITWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
 |||||

QY 61 PRSSQRPVPMGIQHSKELNRTCCINGTCLMGSCFACPPSFYGRNCEHDVYRKENCGSVPH 120  
 |||||

Db 61 PRSSQRPVPMGIQHSKELNRTCCINGTCLMGSCFACPPSFYGRNCEHDVYRKENCGSVPH 120  
 |||||

QY 121 DTWLPKKSCLKCWHQGLRCFPOAFLPGCDGLVMDHLVASRTPPELPPSARTTTFMLVGI 180  
 |||||

Db 121 DTWLPKKSCLKCWHQGLRCFPOAFLPGCDGLVMDHLVASRTPPELPPSARTTTFMLVGI 180  
 |||||

QY 181 CLSIQSYI 188  
 |||||

Db 181 CLSIQSYI 188  
 |||||

RESULT 8  
 ADC78799

ID ADC78799 standard; protein; 188 AA.

AC ADC78799;

DT 01-JAN-2004 (first entry)

DE Human PRO protein #14.

XX human; PRO; cancer; inflammatory bowel disease; ulcerative colitis;  
 KW Crohn's disease.

XX Homo sapiens.

XX WO2003034984-A2.

XX 01-MAY-2003.

XX 15-OCT-2002; 2002WO-US033070.

XX 19-OCT-2001; 2001US-0340083P.

XX (GETH ) GENENTECH INC.

XX Goddard A, Gurney AL;

XX WPI; 2003-481990/45.

DR N-PSDB; ADC78798.

XX New PRO polynucleotide and polypeptide, useful for the manufacture of a  
 PT medicament for diagnosing or treating cancer or inflammatory bowel  
 PT disorder e.g., ulcerative colitis or Crohn's disease.

XX

PS Claim 12; SEQ ID NO 28; 327pp; English.

XX The invention comprises the amino acid and coding sequences of human PRO  
 CC proteins. The DNA and protein sequences of the invention are useful for  
 CC the diagnosis and treatment of cancer and inflammatory bowel disease  
 CC (e.g. ulcerative colitis or Crohn's disease). The present amino acid  
 CC sequence represents a human PRO protein of the invention.

XX Sequence 188 AA;

Query Match 100.0%; Score 1053; DB 7; Length 188;

Best Local Similarity 100.0%; Pred. No. 1.3e-82;

Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCRKMARFSYSVITWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
 |||||

Db 1 MDCRKMARFSYSVITWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWQEEPAIR 60  
 |||||

QY 61 PRSSQRPVPMGIQHSKELNRTCCINGTCLMGSCFACPPSFYGRNCEHDVYRKENCGSVPH 120  
 |||||

Db 61 PRSSQRPVPMGIQHSKELNRTCCINGTCLMGSCFACPPSFYGRNCEHDVYRKENCGSVPH 120  
 |||||

QY 121 DTWLPKKSCLKCWHQGLRCFPOAFLPGCDGLVMDHLVASRTPPELPPSARTTTFMLVGI 180  
 |||||

Db 121 DTWLPKKSCLKCWHQGLRCFPOAFLPGCDGLVMDHLVASRTPPELPPSARTTTFMLVGI 180  
 |||||

QY 181 CLSIQSYI 188  
 |||||

Db 181 CLSIQSYI 188  
 |||||

RESULT 9  
 ADD26058

ID ADD26058 standard; protein; 188 AA.

AC ADD26058;

XX 15-JAN-2004 (first entry)

DT CR-1 cripto protein.

DE Cripto; Cytostatic; tumour; cell proliferation; antibody.

XX Homo sapiens.

XX WO2003083041-A2.

XX 09-OCT-2003.

XX 01-OCT-2002; 2002WO-US031462.

XX 22-MAR-2002; 2002US-0367002P.

XX 24-APR-2002; 2002WO-US011950.

XX (SANI/) SANICOLA-NADEL M.

XX (ADKI/) ADKINS H.

XX (MIKL/) MIKLASZ S D.

XX (RAYH/) RAYHORN P.

XX (SCHL/) SCHIFFER S G.

XX (WILL/) WILLIAMS K.

XX Sanicola-Nadel M, Adkins H, Miklasz SD, Rayhorn P, Schiffer SG;

PI Williams K;

XX WPI; 2003-779452/73.

XX New anti-Cripto antibodies, useful in the therapy of malignant or benign  
 PT tumors of mammals, or for treating a mammal for a condition associated  
 PT with undesired cell proliferation.

XX Example 1; SEQ ID NO 1; 89pp; English.

XX The present invention relates to an antibody that specifically binds to

CC Cripto, to an epitope in the ligand/receptor binding domain of Cripto, or  
 CC to an epitope comprising the domain spanning amino acid residues 46-62.  
 CC The antibodies or compositions comprising them are useful for inhibiting  
 CC growth of tumour cells, treating a mammal having a tumor that  
 CC overexpresses Cripto, or treating a mammal for a condition associated  
 CC with undesired cell proliferation. The antibodies are useful in the  
 CC therapy of malignant or benign tumors of mammals where growth of the  
 CC tumor is at least partially dependent upon Cripto. The present sequence  
 CC represents a CR-2 cripto protein of the invention.  
 XX  
 SQ Sequence 188 AA;

Query Match 100.0%; Score 1053; DB 7; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-82;  
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MDCRMARFSYSVIWIWMAISKVFEGLVAGLGHQEFARPSRGYLAFRDSDSIWQOEPAIR 60  
 Db 1 MDCRMARFSYSVIWIWMAISKVFEGLVAGLGHQEFARPSRGYLAFRDSDSIWQOEPAIR 60  
 Qy 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEHDVRKENGCVPH 120  
 Db 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEHDVRKENGCVPH 120  
 Qy 121 DTWLPKKCSLCKCKWHGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180  
 Db 121 DTWLPKKCSLCKCKWHGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180  
 Qy 181 CLSIQSY 188  
 Db 181 CLSIQSY 188

## RESULT 10

AD182176 ID AD182176 standard; protein; 188 AA.

XX AD182176;

XX 22-APR-2004 (first entry)

XX Human teratocarcinoma-derived growth factor 2.

XX Human: embryonic stem cell; pluripotent stem cell; abnormal cell growth;  
 XX malignancy; differentiation.

XX Homo sapiens.

XX US2003224411-A1.

XX 04-DEC-2003.

XX 13-MAR-2003; 2003US-00388578.

XX 13-MAR-2003; 2003US-00388578.

XX (STAN/) STANTON L W.

XX (BRAN/) BRANDENBERGER R.

XX (GOLD/) GOLD J D.

XX (IRVI/) IRVING J M.

XX (MAND/) MANDALAM R.

XX (MOKM/) MOK M.

XX (SHEL/) SHELTON D.

XX Stanton LW, Brandenberger R, Gold JD, Irving JM, Mandalam R;

PI Mok M, Shelton D;

XX WPI; 2004-119701/12.

DR N-P8DB; AD182175.

XX Assessing culture of undifferentiated primate pluripotent stem cells by  
 PT detecting expression of markers e.g., Zic family member 3, other than  
 PT human telomerase reverse transcriptase/octamer binding transcription

PT factor.

XX Claim 1; SEQ ID NO 6; 106pp; English.

XX The invention relates to assessing a culture of undifferentiated primate  
 CC pluripotent stem cells (pPS, e.g. embryonic stem cells), involving  
 CC detecting expression of markers (MRI) e.g. Zic family member 3 (ZIC3), as  
 CC given in specification, other than human telomerase reverse transcriptase  
 CC (hTERT) or octamer binding transcription factor (Oct)3/4, or a marker  
 CC (MR2) such as cripto or podocalyxin-like protein and hTERT and/or Oct3/4  
 CC or second marker chosen from (MR2). Also included are maintaining (M2)  
 CC pPS cells in a pluripotent state (involves causing them to express one of  
 CC the following markers (MR3) at a higher level, FOXO1A, ZIC3, hypothetical  
 CC protein FLJ20582, Forkhead box H1 (FOXH1), Zinc finger protein, Hsa12,  
 CC KRAB-zinc finger protein SZF1-1 or zinc finger protein of cerebellum  
 CC ZIC2, or any other marker (MR4) chosen from PHD protein Jade-1 (Jade-1),  
 CC kruppel-like zinc finger protein (ZNF300), etc., as given in the  
 CC specification), causing pPS cells to differentiate into a particular  
 CC tissue type by causing them to express one of the markers chosen from  
 CC (MR3) or (MR4) (or markers chosen from GATA binding protein 3 (GATA3),  
 CC core promoter element binding protein (COPBP), etc., as given in the  
 CC specification), maintaining pPS cells in a pluripotent state (involves  
 CC culturing pPS cells or their progeny in the presence of a normally  
 CC secreted protein that is encoded by a gene that down-regulated upon  
 CC differentiation of human embryonic stem (hES) cells, chosen from  
 CC Fibrillin 3 gene, LEFT B gene, ZIC3 gene, BPHAI gene, etc., as given in  
 CC the specification), causing pPS cells to differentiate (involves  
 CC culturing pPS cells or their progeny in the presence of a normally  
 CC secreted protein that is encoded by a gene that up-regulated upon  
 CC differentiation of hES cells, chosen from P311 protein gene, Tax  
 CC interaction protein 1 gene, KIAA0853 protein gene, keratin 19 (KRT 19)  
 CC gene, etc., as given in the specification), causing an encoding sequence  
 CC to be preferentially expressed in undifferentiated pPS cells, causing an  
 CC encoding sequence to be preferentially expressed in differentiated cells,  
 CC sorting (M4) differentiated cells from less differentiated cells  
 CC (involves separating cells expressing a surface marker chosen from any  
 CC one of MRI from cells not expressing the marker), causing pPS cells to  
 CC proliferate without differentiation, identifying genes that are up or  
 CC down regulated during differentiation of pPS cells, and a kit (I) for  
 CC assessing a culture of pPS cells by M1. The method (M1) is useful for  
 CC assessing culture of undifferentiated primate pluripotent stem cells and  
 CC for assessing the growth characteristics of a cell population. The cell  
 CC population has been obtained by culturing cells from human blastocyst or  
 CC from a human patient suspected of having a clinical condition related to  
 CC abnormal cell growth. The method further involves determining whether the  
 CC cell population is pluripotent from the marker expression and assessing  
 CC whether the patient has a malignancy from the marker expression. The  
 CC present sequence is a protein whose expression is down regulated in  
 CC pluripotent stem cells.

XX Sequence 188 AA;

Query Match 100.0%; Score 1053; DB 8; Length 188;

Best Local Similarity 100.0%; Pred. No. 1.3e-82;

Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDCRMARFSYSVIWIWMAISKVFEGLVAGLGHQEFARPSRGYLAFRDSDSIWQOEPAIR 60

Db 1 MDCRMARFSYSVIWIWMAISKVFEGLVAGLGHQEFARPSRGYLAFRDSDSIWQOEPAIR 60

Qy 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEHDVRKENGCVPH 120

Db 61 PRSSQRPVPMGIQHSKELNRTCCCLNGGTCMLGSCFACPPSFYGRNCEHDVRKENGCVPH 120

Qy 121 DTWLPKKCSLCKCKWHGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180

Db 121 DTWLPKKCSLCKCKWHGQLRCFPQAFPLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180

Qy 181 CLSIQSY 188

Db 181 CLSIQSY 188

RESULT 11  
ADR70486  
ID ADR70486 standard; protein; 188 AA.  
XX AC ADR70486;  
XX DT 02-DEC-2004 (first entry)  
XX DE Human teratocarcinoma-derived growth factor (Cripto).  
XX KW Human; telomerase reverse transcriptase; TERT;  
XX KW POU domain, class 5 transcription factor; POU5F1; Oct3; Oct4;  
XX KW teratocarcinoma-derived growth factor; Cripto; podocalyxin-like; PODXL;  
XX KW gastrin-releasing peptide receptor; GRPR; human embryonic stem cell; hES;  
XX KW primate pluripotent stem cell; cancer; gene expression; cell separation;  
XX KW differentiation.  
XX OS Homo sapiens.  
XX PN US2004180347-A1.  
XX PD 16-SEP-2004.  
XX PF 13-MAR-2003; 2003US-00389431.  
XX PR 13-MAR-2003; 2003US-00389431.  
XX PA (STAN/) STANTON L W.  
XX PA (BRAN/) BRANDENBERGER R.  
XX PA (GOLD/) GOLD J D.  
XX PA (IRVI/) IRVING J M.  
XX PA (MAND/) MANDALAM R.  
XX PA (MOKM/) MOK M.  
XX PI Stanton LW, Brandenberger R, Gold JD, Irving JM, Mandalam R;  
PI Mok M;  
XX WPI: 2004-675599/66.  
XX DR N-PSDB; ADR70485.  
XX DR Assessing culture of undifferentiated human embryonic stem cells or their  
XX PT progeny, by detecting Cripto, gastrin-releasing peptide (GRP) receptor  
XX PT and podocalyxin-like protein markers, and either hTERT and/or Oct3/4, or  
XX PT GRP receptor.  
XX PS Disclosure; SEQ ID NO 6; 57pp; English.  
XX CC The invention relates to assessing a culture of undifferentiated human  
XX CC embryonic stem (hES) cells (undifferentiated primate pluripotent stem  
XX CC cells) or their progeny, involves detecting or measuring a marker such as  
XX CC Cripto (teratocarcinoma-derived growth factor), gastrin-releasing peptide  
XX CC (GRP) receptor and podocalyxin-like protein, and either hTERT (telomerase  
XX CC reverse transcriptase) and/or Oct3/4 (also known as POU domain, class 5,  
XX CC transcription factor 1 (POU5F1)), or GRP receptor. The method involves  
XX CC detecting or measuring at least two markers, and detecting or measuring  
XX CC hTERT and/or Oct3/4. The expression of the marker(s) is detected or  
XX CC measured at mRNA level by PCR amplification. The expression of the  
XX CC marker(s) is detected or measured at the protein level by antibody assay.  
XX CC The method involves quantifying the proportion of undifferentiated hES  
XX CC cells or differentiated cells in the culture from the marker expression.  
XX CC The level of the marker is determined to be at least 100-fold higher than  
XX CC the level of the marker in BJ fibroblasts or is determined to be no less  
XX CC than 100-fold lower than the level of the marker in hES cells, cultured  
XX CC on an extracellular matrix in medium conditioned with mouse embryonic  
XX CC fibroblasts and containing 4 ng/ml basic fibroblast growth factor. The  
XX CC method further involves modifying the culture conditions so as to cause  
XX CC the hES cells to increase expression of the marker detected or measured  
XX CC in the culture. The method is useful for assessing a culture of  
XX CC undifferentiated hES cells or their progeny. The marker used in the above  
XX CC method is useful for characterising pluripotent stem cells and their  
XX CC differentiated progeny, for clinical diagnosis of cancer, for assessing  
XX CC and manipulating culture conditions, regulating gene expression, cell  
XX CC separation and purification, and to influence differentiation. The

CC present sequence is a marker protein of the invention for  
CC undifferentiated stem cells.  
XX SQ Sequence 188 AA;  
Query Match 100.0%; Score 1053; DB 8; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.3e-82;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDCRKMARFSYSVITWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60  
DB 1 MDCRKMARFSYSVITWIMAIKVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEEPAIR 60  
QY 61 PRSSQVPPMGIQHSKELNRTCLNGTCLMGSCFACPPSFYGRNCHEHDYKENCVSVPH 120  
DB 61 PRSSQVPPMGIQHSKELNRTCLNGTCLMGSCFACPPSFYGRNCHEHDYKENCVSVPH 120  
QY 121 DTWLPKCKSLCKWHGQLRCFPOAFLPGCDGLVMDHVLVARTPELPSPARTTTFMLVGI 180  
DB 121 DTWLPKCKSLCKWHGQLRCFPOAFLPGCDGLVMDHVLVARTPELPSPARTTTFMLVGI 180  
QY 181 CLSIQSYV 188  
DB 181 CLSIQSYV 188  
RESULT 12  
ADR79200  
ID ADR79200 standard; protein; 188 AA.  
XX AC ADR79200;  
XX DT 30-DEC-2004 (first entry)  
XX DE Human Cripto protein.  
XX KW Cripto protein; vaccine; cancer; lung; colon; colorectal; breast; T cell;  
XX KW immune response; cytostatic; human; gene; ds; variant; mutant.  
XX OS Homo sapiens.  
XX PN US2004202648-A1.  
XX PD 14-OCT-2004.  
XX PF 01-APR-2004; 2004US-00816476.  
XX PR 04-APR-2003; 2003US-0460733P.  
XX PA (CABE/) CABEZON T E V S.  
XX PA (GERA/) GERARD C M G.  
XX PA (PALM/) PALMANTIER R M.  
XX PA (VINA/) VINALS Y D B C.  
XX PI Cabezon TEVS, Gerard CMG, Palmantier RM, Vinals YDBC;  
XX WPI: 2004-727978/71.  
XX DR N-PSDB; ADR79199.  
XX PT Novel immunogenic fragment of Cripto polypeptide, useful as vaccine for  
XX PT treating cancer e.g. colon, lung, colorectal and breast cancer.  
XX PS Example 2; SEQ ID NO 6; 64pp; English.  
XX CC The invention relates to immunogenic fragments of Cripto polypeptide  
XX CC which are used as vaccines. The immunogenic fragment of Cripto protein is  
XX CC useful for treating or diagnosing cancer preferably lung, colon,  
XX CC colorectal or breast cancer, for stimulating T cells specific for Cripto  
XX CC and for inhibiting the development of a cancer in a patient. It is useful  
XX CC for inducing an immune response to Cripto in an animal. The present  
XX CC sequence is the human Cripto protein.  
XX SQ Sequence 188 AA;

Query Match	100.0%;	Score 1053;	DB 8;	Length 188;	
Best Local Similarity	100.0%;	Pred. No. 1.3e-82;			
Matches 188;	Conservative	0;	Mismatches	0;	Indels
				0;	Gaps
					0;
Qy	1	MDCRKFYSYVIWIMAI	SKVFLVAGLGHQEFARPSRGYLA	FRDDSIWPOEPAIR	60
Db	1	MDCRKFYSYVIWIMAI	SKVFLVAGLGHQEFARPSRGYLA	FRDDSIWPOEPAIR	60
Qy	61	PRSSQRPVPMGIQHSKELNRTCC	LMGTCMLGSCACPPSPFYGRNC	CHDVVKENCGSVPH	120
Db	61	PRSSQRPVPMGIQHSKELNRTCC	LMGTCMLGSCACPPSPFYGRNC	CHDVVKENCGSVPH	120
Qy	121	DTWLPKCSLCKWHGQLRCFPQAF	PGCDGLVMDHLVASRTPELPPS	ARTTTFMLVGI	180
Db	121	DTWLPKCSLCKWHGQLRCFPQAF	PGCDGLVMDHLVASRTPELPPS	ARTTTFMLVGI	180
Qy	181	CLSIQSYI	188		
Db	181	CLSIQSYI	188		
RESULT 14					
ADY85961					
ID	ADY85961	standard;	protein;	188	AA.
XX					
AC	ADY85961;				
XX					
DT	16-JUN-2005	(first entry)			
XX					
DE	Human Cripto oncoprotein Seq 2.				
XX					
KW	cell signaling; oncogenesis; antisense therapy; cytostatic; Cripto;				
KW	mutagenesis; protein interaction; oncoprotein.				
XX					
OS	Homo sapiens.				
XX					
FH	Key	Location/Qualifiers			
FT	Modified-site	88			
FT		/note= "Modified by O-fucosylation"			
PN	WO2005028433-A2.				
XX					
PD	31-MAR-2005.				
XX					
PF	14-SEP-2004; 2004WO-US029967.				
XX					
PR	15-SEP-2003; 2003US-0503046P.				
XX					
PA	(RERE-) RES DEV FOUND.				
XX					
PI	Vale W, Gray PC, Harrison CA;				
XX					
DR	WPI; 2005-242562/25.				
XX					
PT	Augmenting signaling of a ligand of receptor serine kinase in a cell				
PT	comprises inhibiting the formation of complexes between Cripto and the				
PT	ligand on the surface of the cell.				
XX					
PS	Disclosure; SEQ ID NO 2; 60pp; English.				
XX					
CC	This invention relates to a novel method for augmenting signaling of a				
CC	ligand of a receptor serine kinase within a cell by inhibiting the				
CC	formation of complexes between Cripto and this ligand on the surface of				
CC	the cell. Specifically, it refers to TGF-beta and activin which are the				
CC	ligands of serine kinase receptors and which regulate tissue homeostasis				
CC	by activating the Smad2/3 intracellular signaling pathway; disruption of				
CC	this signaling pathway is associated with oncogenesis and tumorigenesis.				
CC	As such, the present invention describes a method for augmenting Smad2/3				
CC	signaling in a cell by administering a mutant ligand that retains				
CC	signaling activity but is unable to bind to Cripto, and thus bypasses				
CC	antagonism by Cripto. Note that augmentation of signaling increases				
CC	phosphorylation and activation of Smad2 and Smad3 in the cell, such that				
CC	it decreases the proliferative rate of the cell. The receptor serine				
CC	kinase is a type I activin receptor-like kinases-4 or -5 (ALK-4 or ALK-5)				
CC	and the formation of complexes is inhibited by suppressing expression of				
CC	Cripto using antisense oligonucleotides (siRNA) directed against Cripto,				

Query Match	100.0%;	Score 1053;	DB 8;	Length 188;	
Best Local Similarity	100.0%;	Pred. No. 1.3e-82;			
Matches 188;	Conservative	0;	Mismatches	0;	Indels
				0;	Gaps
					0;
Qy	1	MDCRKFYSYVIWIMAI	SKVFLVAGLGHQEFARPSRGYLA	FRDDSIWPOEPAIR	60
Db	1	MDCRKFYSYVIWIMAI	SKVFLVAGLGHQEFARPSRGYLA	FRDDSIWPOEPAIR	60
Qy	61	PRSSQRPVPMGIQHSKELNRTCC	LMGTCMLGSCACPPSPFYGRNC	CHDVVKENCGSVPH	120
Db	61	PRSSQRPVPMGIQHSKELNRTCC	LMGTCMLGSCACPPSPFYGRNC	CHDVVKENCGSVPH	120
Qy	121	DTWLPKCSLCKWHGQLRCFPQAF	PGCDGLVMDHLVASRTPELPPS	ARTTTFMLVGI	180
Db	121	DTWLPKCSLCKWHGQLRCFPQAF	PGCDGLVMDHLVASRTPELPPS	ARTTTFMLVGI	180
Qy	181	CLSIQSYI	188		
Db	181	CLSIQSYI	188		
RESULT 13					
ADT79197					
ID	ADT79197	standard;	protein;	188	AA.
XX					
AC	ADT79197;				
XX					
DT	30-DEC-2004	(first entry)			
XX					
DE	Human Cripto 1 protein.				
XX					
KW	Cripto protein; vaccine; cancer; lung; colon; colorectal; breast; T cell;				
KW	immune response; cytostatic; human.				
XX					
OS	Homo sapiens.				
XX					
PN	US2004202648-A1.				
XX					
PD	14-OCT-2004.				
XX					
PF	01-APR-2004; 2004US-00816476.				
XX					
PR	04-APR-2003; 2003US-0460733P.				
XX					
PA	(CABE/) CABEZON T E V S.				
PA	(GERA/) GERARD C M G.				
PA	(PALM/) PALMANTIER R M.				
PA	(VINA/) VINALS Y D B C.				
XX					
PI	Cabezon TEVS, Gerard CMG, Palmantier RM, Vinals YDBC;				
XX					
DR	WPI; 2004-727978/71.				
DR	N-PSDB; ADT79195.				
XX					
PT	Novel immunogenic fragment of Cripto polypeptide, useful as vaccine for				
PT	treating cancer e.g. colon, lung, colorectal and breast cancer.				
XX					
PS	Claim 1; SEQ ID NO 3; 64pp; English.				
XX					
CC	The invention relates to immunogenic fragments of Cripto polypeptide				
CC	which are used as vaccines. The immunogenic fragment of Cripto protein is				
CC	useful for treating or diagnosing cancer preferably lung, colon,				
CC	colorectal or breast cancer, for stimulating T cells specific for Cripto				
CC	and for inhibiting the development of a cancer in a patient. It is useful				
CC	for inducing an immune response to Cripto in an animal. The present				
CC	sequence is the human Cripto 1 protein.				
XX					
SQ	Sequence 188 AA;				
Query Match	100.0%;	Score 1053;	DB 8;	Length 188;	
Best Local Similarity	100.0%;	Pred. No. 1.3e-82;			
Matches 188;	Conservative	0;	Mismatches	0;	Indels
				0;	Gaps
					0;

CC and also mutating at least one allele of Cripto by homologous  
CC recombination. Accordingly, pharmaceutical compositions derived thereof  
CC exhibit cytostatic activity. This polypeptide sequence is the human  
CC Cripto oncoprotein of the invention.  
XX  
SQ Sequence 188 AA;  
  
Query Match 100.0%; Score 1053; DB 9; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.3e-82;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MDCRKMARFSYSVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEPAIR 60  
Db 1 MDCRKMARFSYSVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEPAIR 60  
  
Qy 61 PRSSQRVPPMGIQHSKELNRTCCCLNGTCLMGSGFCACPPSFYGRNCHDVKENCSSVPH 120  
Db 61 PRSSQRVPPMGIQHSKELNRTCCCLNGTCLMGSGFCACPPSFYGRNCHDVKENCSSVPH 120  
  
Qy 121 DTWLPKKCSLCKWHGQLRCFPOAFLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180  
Db 121 DTWLPKKCSLCKWHGQLRCFPOAFLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180  
  
Qy 181 CLSIQSYI 188  
Db 181 CLSIQSYI 188  
  
RESULT 15  
AD242238  
ID AD242238 standard; protein; 188 AA.  
XX AC AD242238;  
XX AC  
XX DT 30-JUN-2005 (first entry)  
XX DE Human Cripto-1 protein.  
XX KW neurodegenerative disease; neuroprotective; gene therapy.  
XX OS Homo sapiens.  
XX PN WO2005033341-A2.  
XX PD 14-APR-2005.  
XX PF 01-OCT-2004; 2004WO-US032649.  
XX PR 03-OCT-2003; 2003US-0508750P.  
XX PA (UNIV ) UNIV KANSAS MEDICAL CENT.  
XX PA (USSH ) US DEPT OF HEALTH & HUMAN SERVICES.  
XX PI Salomon D, Berman N, Stephens E;  
XX DR WPI; 2005-273509/28.  
XX DR N-PSDB; AD242237.  
XX PT Detecting a neurodegenerative disease, such as NeuroAIDS, Alzheimer's  
XX PT disease, multiple sclerosis, Parkinson's disease and encephalitis, by  
XX PT assaying the copy number of a Cripto-1 gene or the expression level of a  
XX PT Cripto-1 gene product.  
XX PS Disclosure; SEQ ID NO 2; 42pp; English.  
XX  
CC The invention relates to a novel method for detecting a neurodegenerative  
CC disease in a mammal. The method comprises assaying the copy number of a  
CC Cripto-1 gene (also known as Teratocarcinoma-derived Growth Factor-1  
CC (TDGF-1) gene) or the expression level of a Cripto-1 gene product in the  
CC central nervous system of the mammal, where an amplification of the  
CC Cripto-1 gene or an overexpression of the Cripto-1 gene product is  
CC indicative of a neurodegenerative disease in the mammal. The invention  
CC further comprises: a method for inhibiting progression of a

CC neurodegenerative disease in a mammal, comprising administering to the  
CC mammal an agent that inhibits Cripto-1 in the central nervous system of  
CC the mammal, where the progression of the neurodegenerative disease is  
CC inhibited; and an isolated or purified oligonucleotide consisting  
CC essentially of SEQ ID NO: 3 (AD242239) or 4 (AD242240). The agent may be  
CC used in the treatment of neurodegenerative disorders by gene therapy. The  
CC methods and compositions of the present invention are useful for the  
CC diagnosis, prevention and/or treatment of neurodegenerative disorders,  
CC such as NeuroAIDS, Alzheimer's disease, multiple sclerosis, amyotrophic  
CC lateral sclerosis (ALS), Parkinson's disease, and encephalitis. This  
CC sequence represents the human Cripto-1 protein of the invention.  
XX  
SQ Sequence 188 AA;

Query Match 100.0%; Score 1053; DB 9; Length 188;  
Best Local Similarity 100.0%; Pred. No. 1.3e-82;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MDCRKMARFSYSVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEPAIR 60  
Db 1 MDCRKMARFSYSVIWIMAIKSVFELGLVAGLGHQEFARPSRGYLAFRDDSIWPOEPAIR 60  
  
Qy 61 PRSSQRVPPMGIQHSKELNRTCCCLNGTCLMGSGFCACPPSFYGRNCHDVKENCSSVPH 120  
Db 61 PRSSQRVPPMGIQHSKELNRTCCCLNGTCLMGSGFCACPPSFYGRNCHDVKENCSSVPH 120  
  
Qy 121 DTWLPKKCSLCKWHGQLRCFPOAFLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180  
Db 121 DTWLPKKCSLCKWHGQLRCFPOAFLPGCDGLVMDHLVASRTPELPSPARTTTFMLVGI 180  
  
Qy 181 CLSIQSYI 188  
Db 181 CLSIQSYI 188

Search completed: February 3, 2006, 18:22:37  
Job time : 178.72 secs

**THIS PAGE BLANK (USPTO)**